CONFORMATIONAL ASPECTS OF PROLINE HYDROXYLATION IN COLLAGEN BIOSYNTHESIS STUDIES WITH SYNTHETIC PEPTIDES

PRABHAKARA LAKSHMI ATREYA

Conformational Aspects of Proline Hydroxylation in Collagen Biosynthesis

Studies with Synthetic Peptides

BY

© **Prabhakara Lakshmi Atreya**, B.Sc., M.Sc.

A thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Department of Biochemistry

Memorial University of Newfoundland

October 1987

St. John's Newfoundland

ISBN 0-315-50501-X

SUBJECT CATEGORIES

Dissertation Abstracts International is arranged by broad, general subject categories. Choose the *one* listed below (capital letters) which most nearly describes the general content of your dissertation. If the major subject category has sub-fields under it, and *only* if it does, please choose *one* (small letters). (Ex.: ECONOMICS, Theory). Enter subject category on Item 7 of Agreement Form.

HUMANITIES

IA COMMUNICATIONS AND THE ARTS
- ARCHITECTURE
- CINEMA
- FINE ARTS
- INFORMATION SCIENCE
- JOURNALISM
- LIBRARY SCIENCE
- MASS COMMUNICATIONS
- MUSIC
- SPEECH
- THEATER

IIA EDUCATION
- EDUCATION
  - General
  - Administration
  - Adult
  - Agricultural
  - Art
  - Audiovisual
  - Business
  - Community and Social
  - Community Colleges
  - Curriculum and Instruction
  - Early Childhood
  - Elementary
  - Finance
  - Guidance and Counseling
  - Health
  - Higher
  - History
  - Home Economics
  - Industrial
  - Language and Languages
  - Mathematics
  - Middle School
  - Minorities
  - Music
  - Personality Development and Mental Hygiene
  - Philosophy
  - Physical
  - Preschool
  - Programmed Instruction
  - Psychology
  - Religion
  - Sciences
  - Secondary
  - Social Sciences
  - Special
  - Teacher Training
  - Teaching Machines
  - Tests and Measurements
  - Theory and Practice
  - Vocational

IIIA LANGUAGE, LITERATURE AND LINGUISTICS
- LANGUAGE
  - General
  - Ancient
  - Linguistics
  - Modern
- LITERATURE
  - General
  - Classical
  - Comparative
  - Medieval
  - Modern
  - American
  - Asian
  - Dutch and Scandinavian
  - English
  - Germanic
  - Latin American
  - Romance
  - Russian and East European
  - Slavic and Finno-Ugrian

IVA PHILOSOPHY, RELIGION AND THEOLOGY
- PHILOSOPHY
- RELIGION
  - General
  - Clergy
  - History
  - Music
  - Philosophy
- THEOLOGY

VA SOCIAL SCIENCES
- ACCOUNTING
- AMERICAN STUDIES
- ANTHROPOLOGY
  - Archaeology
  - Cultural
  - Physical
- BANKING
- BUSINESS ADMINISTRATION
- ECONOMICS
  - General
  - Agricultural
  - Commerce-Business
  - Finance (includes Public Finance)
  - History
  - Theory
- FOLKLORE
- HISTORY
  - General
  - Ancient
  - Medieval
  - Modern
  - Black
  - Church
  - Africa
  - Asia
  - Australia and Oceania
  - Canada
  - Europe
  - Latin America
  - United States
- HISTORY OF SCIENCE
- LAW
- MANAGEMENT
- MARKETING
- POLITICAL SCIENCE
  - General
  - International Law and Relations
  - Public Administration
- PUBLIC RELATIONS
- RECREATION
- SOCIAL GEOGRAPHY
- SOCIAL STRUCTURE
- SOCIAL WORK
- SOCIOLOGY
  - General
  - Community Organization
  - Criminology
  - Demography
  - Educational
  - Individual and Family Studies
  - Industrial
  - Labor Relations
  - Public Welfare
  - Race Relations
  - Social Problems
  - Statistics—Research Methods
  - Theory
- TRANSPORTATION
- URBAN AND REGIONAL PLANNING
- WOMEN'S STUDIES

SCIENCES

IB BIOLOGICAL SCIENCES
- AGRICULTURE
  - General
  - Animal Culture
  - Animal Pathology
  - Forestry & Wildlife
  - Plant Culture
  - Plant Pathology
  - Plant Physiology
  - Range Management
  - Wood Technology
- AGRONOMY
- ANATOMY
- BIOLOGICAL OCEANOGRAPHY
- BIOLOGY
- BIOPHYSICS
  - General
  - Medical
- BIOSTATISTICS
- BOTANY
- ECOLOGY
- ENTOMOLOGY
- GENETICS
- LIMNOLOGY
- MICROBIOLOGY
- PHYSIOLOGY
- RADIATION BIOLOGY
- VETERINARY SCIENCE
- ZOOLOGY

IIB EARTH SCIENCES
- GEOCHEMISTRY
- GEODESY
- GEOLOGY
- GEOPHYSICS
- HYDROLOGY
- MINERALOGY
- PALEOBOTANY
- PALEONTOLOGY
- PALEOZOOLOGY
- PHYSICAL GEOGRAPHY
- PHYSICAL OCEANOGRAPHY

IIIB HEALTH AND ENVIRONMENTAL SCIENCES
- ENVIRONMENTAL SCIENCES
- FOOD TECHNOLOGY
- HEALTH SCIENCES
  - General
  - Audiology
  - Chemotherapy
  - Dentistry
  - Education
  - Hospital Management
  - Human Development
  - Hygiene
  - Immunology
  - Medicine & Surgery
  - Mental Health
  - Nursing
  - Nutrition
  - Pathology
  - Pharmacy
  - Public Health
  - Radiology
  - Recreation
  - Speech Pathology
- HOME ECONOMICS
- PHARMACOLOGY

IVB PHYSICAL SCIENCES

PURE SCIENCES
- CHEMISTRY
  - General
  - Analytical
  - Biological
  - Inorganic
  - Nuclear
  - Organic
  - Pharmaceutical
  - Physical
  - Polymer
  - Radiation
  - Water
- MATHEMATICS
- PHYSICS
  - General
  - Acoustics
  - Astronomy & Astrophysics
  - Atmospheric Science
  - Atomic
  - Electronics and Electricity
  - Elementary Particles and High Energy
  - Fluid and Plasma
  - Molecular
  - Nuclear
  - Optics
  - Radiation
  - Solid State
- STATISTICS

APPLIED SCIENCES
- APPLIED MECHANICS
- ASTRONAUTICS
- COMPUTER SCIENCE
- ENGINEERING
  - General
  - Aeronautical
  - Agricultural
  - Automotive
  - Biomedical
  - Chemical (includes ceramics and fuel)
  - Civil
  - Electronics and Electrical
  - Heat and Thermodynamics
  - Hydraulic
  - Industrial
  - Marine
  - Materials Science
  - Mechanical
  - Metallurgy
  - Mining
  - Nuclear
  - Petroleum
  - Sanitary and Municipal
  - System Science
- OPERATIONS RESEARCH
- PLASTICS TECHNOLOGY
- TEXTILE TECHNOLOGY

VB PSYCHOLOGY
- PSYCHOLOGY
  - General
  - Clinical
  - Experimental
  - Industrial
  - Physiological
  - Psychobiology
  - Social

Abstract

The hydroxylation of selected proline residues by prolylhydroxylase (E.C. 1.14.11.2) is a crucial posttranslational event in the biosynthesis of collagen, an important protein of the connective tissues. Hydroxyproline (Hyp) offers additional stability to the unique triple-helical conformation of collagen, which in turn, is necessary for the functional viability of the protein, at physiological temperatures.

Earlier studies on the substrate specificity of prolylhydroxylase have been intriguing. It was proposed earlier that prolylhydroxylase recognizes the folded β-turn conformation, formed at the Pro-Gly segments in the nascent procollagen chains (Brahmachari and Ananthanarayanan, 1979). The present thesis involves the further elucidation of conformational aspects of proline hydroxylation *in vitro*, using chicken prolylhydroxylase and Pro-containing synthetic peptides.

Pure prolylhydroxylase was obtained from 13-day old chicken embryos using established procedures. Pro-containing linear oligopeptides were characterized in different solvents, using circular dichroism (CD) and infrared (IR) spectroscopy. These studies have indicated the existence of two conformations, namely, an extended conformation similar to that of poly(Pro) (PP-II) and a folded β-turn, in these peptides. The interaction between the enzyme and the oligopeptides of

known conformation was studied by the following reactions: (1) hydroxylation of the peptides themselves and (2) the capability of these peptides to compete with the standard substrate, for the active site of prolylhydroxylase. It was found that peptides with either β-turn or extended conformation alone can act only as inhibitors. On the other hand, peptides with both these conformations can also serve as substrates for the enzyme, in addition to being competitive inhibitors.

Based on these observations, a model is proposed for the conformational criteria of enzymatic proline hydroxylation. According to this model, the enzyme requires the presence of PP-II like extended conformation followed by folded β-turns in the substrate molecules. The PP-II structure is necessary at the binding site of the enzyme, while the β-turn structure is necessary at the catalytic site. Peptides with either one of these structures can act only as inhibitors since they can fulfill only part of the conformational requirement. These studies are of importance, since they help to define the observed substrate specificity of prolylhydroxylase, in precise conformational terms.

The structure-function relationship of the prolylhydroxylase itself and its interaction with substrates and cosubstrates, in conformational terms, are also studied by CD and fluorescence spectroscopy. The implications of these studies in understanding the substrate specificity of prolylhydroxylase are discussed.

Acknowledgements

I would like to thank Memorial University for giving me an opportunity to pursue graduate studies and also for financial support in the form of a fellowship.

I wish to express my sincere gratitude and appreciation to Dr. V.S. Ananthanarayanan for his continued enthusiasm and supervision throughout the course of this work and also for the financial support. Special thanks go to him for critically reading my thesis and several interesting discussions. I also like to thank him for making the computer facilities available during the preparation of the thesis.

I am very grateful to Drs. M.E. Brosnan, T.R. Patel and W.S. Davidson for being active members of my supervisory committee and for their helpful support and encouragement throughout my programme. Special thanks to Dr. M. E. Brosnan for critically reading my thesis and for her constructive criticism and many helpful discussions. I would like to thank Dr. Davidson for his critical perusal of this thesis. I would like to extend my sincere appreciation to Drs. K.M.W. Keough and J.T. Brosnan for their friendly support and for generously letting me use their equipment whenever needed. I would also like to thank Dr. A.D. Rahimtula for the scintillation counter facility.

I would like to thank the other members of our laboratory especially, Dr.

Sam Attah-Poku, Mr. Lorne Taylor, Ms. Lynn Zijenah and Ms. Susan Thompson for creating a very pleasant and interesting working atmosphere in the laboratory. Special thanks to Sam for making available the "best" peptide preparations, without which this work would not have been possible.

I would like to express my sincere thanks to all members of the Biochemistry department, especially the staff in the general office and supply stores whose help has been valuable. I would like to thank Mr. D. Hall for the amino acid analysis. Also the friendly and helpful nature of the staff in the School of Graduate Studies is greatly appreciated. I also thank the staff in the computer services both in the operations and consultation sections, for their help with the computers during the preparation of this thesis.

Last but not the least, I thank my parents for always respecting the freedom of my thoughts and deeds.

Dedicated to
my best friend and
most wonderful husband
for his countless sacrifices
during the course of this study

Table of Contents

List of Tables

List of Figures

Abbreviations

Aib	α-aminobutyrate
α-KG	α-ketoglutarate
β-Ala	β-Alanine
BSA	Bovine Serum Albumin
CBZ	Carbobenzoxy
CD	Circular Dichroism
CNBr	Cyanogen Bromide
C-terminal	Carboxy-terminal
DA	D-Ala
D-Ala	D-Alanine
D-Phe	D-Phenylalanine
DEAE	Diethylaminoethyl
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic Acid
Fe^{+2}	Ferrous Ions
$FeSO_4$	Ferrous Sulphate
FT-IR	Fourier Transform Infrared
HPLC	High Performance Liquid Chromatography
Hyp	4-Hydroxy-L-Proline
Hyl	5-Hydroxy-L-Lysine
Ig	Immunoglobulin
IR	Infrared
MeOH	Methanol
M_r	Relative Molecular Weight
MRW	Mean Residue Weight
NHMe	$NHCH_3$
NMR	Nuclear Magnetic Resonance
N-terminal	Amino-terminal
OBu	O-butyl
OMe	O-methyl
OSt	O-stearyl
Orn	Ornithine
Psi	Pounds Per Square Inch
Piv	Pivaloyl
PAGE	Polyacrylamide Gel Electrophoresis

PP I	Poly(Proline) or poly(Pro) I
PP II	Poly(Proline) or poly(Pro) II
$(P\text{-}P\text{-}G)_5$	$(Pro\text{-}Pro\text{-}Gly)_5$
Sar	Sarcosine, N-methyl Glycine
SDS	Sodium Dodecyl Sulphate
t-Boc	*tertiary*-Butyloxy Carbonyl
TCA	Trichloroacetic Acid
TFE	Trifluoroethanol
TEMED	N,N,N',N'- tetramethylethylenediamine
T_m	Thermal Melting Point
UV	Ultra-Violet

Amino acids	3-letter Code	One-letter code
Alanine	Ala	A
Arginine	Arg	R
Aspartate	Asp	D
Asparagine	Asn	N
Cysteine	Cys	C
Glycine	Gly	G
Glutamate	Glu	E
Glutamine	Gln	Q
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Chapter 1

Introduction

1.1. Introduction

The structure and function of a protein are interdependent in carrying out the processes of life. This phenomenon has been found to be especially true in the case of collagens, a family of proteins ubiquitous in all multicellular organisms. The term collagen is now applied to a series of related, yet genetically distinct, macromolecular species found as major constituents of bones, cartilage, skin, ligaments, cornea, vitreous, eye lens, blood vessels and glandular ducts. They are involved in carrying out such diverse functions as weight bearing, force transmission, mechanical support, matrix formation and cell-cell interactions. The capacity to form supramolecular aggregates in extracellular spaces is one of the properties characteristic of molecules belonging to the collagen family. Collagen's distinctive structural and functional properties have been ascribed to its unique physicochemical characteristics such as amino acid composition, sequence, and conformation. These, in turn, govern the precise organization of collagen fibres, the most highly characterized form of collagen aggregates. This organization of particular collagen fibres in association with other structural molecules of the extracellular matrix appears to determine the final architecture and function of a particular tissue.

The innumerable aspects of collagen's structure and function have resulted in a flood of collagen-related studies in the past several years. The vast collagen-related literature includes a number of monographs and texts (See, for example, Ramachandran 1967; Balazs, 1970; Ramachandran and Reddi, 1976; Fleischmajer, Olsen and Kuhn, 1985). Several comprehensive reviews have also appeared dealing with more 'specific aspects' like collagen chemistry (Kulonen and Pikkarainen, 1970; Fietzek and Kuhn, 1976; Piez, 1976), biosynthesis (Grant and Prockop, 1972; Bornstein, 1974; Kivirikko and Myllyla, 1984), proline hydroxylation (Cardinale and Udenfriend 1974; Kivirikko and Myllyla, 1980; Hanauski-Abel and Gunzler, 1982; Kivirikko and Myllyla, 1985), collagen pathology (Nimni and Deshmukh, 1973; Lapiere and Nusgens, 1976) and collagen genes (see Fleischmajer, Olsen and Kuhn, 1985).

In the following sections, our current knowledge about collagen structure in terms of primary structure and conformation, biosynthesis and posttranslational modifications with special reference to proline hydroxylation will be briefly discussed followed by the objectives and scope of the present thesis.

1.2. Collagen Structure

1.2.1. Collagen Types

Currently, about 13 types of collagen have been identified in vertebrate tissues and characterized to varying degrees (Miller, 1985). In broad terms, the collagen types can be divided according to their extracellular location and the type of supramolecular aggregates formed. Two such major groups exist: (1) fibrillar and (2) pericellular. The first group, called interstitial collagens, are

responsible for the extracellular fabric of the major connective tissues while the second group are of finer texture and predominantly occur in the basement membranes. The basic characteristic feature of all types of collagen molecules is the presence of three polypeptide chains called the α-chains. Each of these polypeptide chains is derived from a distinct genetic locus in the vertebrate genome. Three such chains aggregate to form different types of collagen. Bornstein and Traub (1979) classified collagens into four types based on chemical composition and tissue distribution. Miller (1985) classified different collagens into three groups based on the length, molecular weight and fibril-forming nature of the polypeptide chains. Group 1 contains collagen types I, II, III and V. In general, the chains are made up of continuous helical domains of about 300 nm long and a relative molecular weight (M_r) of about 95,000. Type I has been found in skin, bone, tendon, ligament, fascia, dentin and interstitial connective tissues. Type II has been found predominantly in cartilaginous tissues. Type III collagen is a major constituent of fetal skin, blood vessels and gastro-intestinal tract. Fibroblasts, cartilage and human amniotic membranes have been found to contain type V collagen. Group 2 includes collagen types IV, VI, VII and VIII whose polypeptide chains contain triple-helical domains interspersed by non-helical segments. The chains have a M_r of about 95,000 and form different kinds of fibrils, compared to group 1 molecules. Of these, only type IV has been well-characterized and is a major matrix forming collagen of the basement membranes from kidney glomeruli, lens capsule and human placenta. Group 3 includes types IX and X found in hyaline cartilage and available data suggest that they contain more non-helical than helical domains and therefore, are not capable of forming fibrils (Miller, 1985).

1.2.2. Amino Acid Composition

Collagens from different sources show a highly characteristic amino acid composition (Eastoe, 1967; Fraser and McRae, 1973; Bornstein and Traub, 1979). Glycine is found to be present to an extent of 33% which is very high when compared to a typical globular protein where it is about 5-10%. Another notable feature is the high content of the imino acids proline (Pro) and hydroxyproline which constitute about 20-25 mole% of total amino acids. Hydroxyproline mainly occurs as 4-hydroxy-L-proline and constitutes about 12 mole% of amino acids. 4-hydroxy-L-proline is abbreviated, for convenience, as Hyp throughout this thesis. Recently, Hyp has been found in other proteins like elastin (Gray *et al.*, 1973; Sandberg, 1976), the collagen-like tail structure of C1q subcomponent of human complement (Muller-Eberhard, 1975; Porter and Reid, 1978), acetylcholine esterase (Anglister *et al.*, 1976; Lwebuga-Mukasa *et al.*, 1976; Rosenberg and Richardson, 1977), a lung surfactant protein (Drickamer *et al.*, 1986) and Volvox (Schlifpenbacher *et al.*, 1986). Type IV collagens from basement membranes are found to contain 3-hydroxy-L-proline (3-Hyp) (Kefalides, 1975; Burgeson *et al.*, 1976). 3-Hyp has not been reported in any other proteins except in collagens. Another important amino acid unique to collagen is 5-hydroxy-L-lysine (abbreviated as Hyl). Its presence has also been demonstrated in C1q and in acetylcholine esterase but not in elastin (Kivirikko and Myllyla, 1980). There appear to be only relatively small differences in Hyp content between various types of collagens (Kivirikko and Myllyla, 1980). Also the Hyp content seems to vary only within narrow limits under normal and abnormal conditions (Kivirikko and Myllyla, 1980), whereas a marked variation is found in the amounts of 3-Hyp

and 5-Hyl, not only among different collagen types but also within the same collagen type under different physiological and pathological conditions (Tryggvason *et al.*, 1978; Kivirikko and Myllyla, 1980). Collagens are glycoproteins. Glucosylgalactose and galactose have been found attached to Hyl through O-glycosidic bonds. Enzymes transferring these carbohydrates have been studied (Butler and Cunningham, 1966; Spiro and Spiro 1971, Kivirikko and Myllyla, 1979).

1.2.3. Primary Structure of Collagen

As mentioned earlier (section 1.2.1), the basic collagen molecule common to all types of collagen is composed of three α-chains, each about 1000 residues long. The complete covalent structures of the homologous chains, α1(I), α2(I), α1(III) are known (Piez, 1976; Hoffmann *et al.*, 1980). The α-chains are initially synthesized as the precursor polypeptide chains, pro-α-chains which are about 40% larger than the α-chains of collagen. The procollagen differs from collagen in that it contains additional peptide extensions on both N-terminal and C-terminal ends of the constituent pro-α-chains (Bornstein,1974). The schematic representation of type I procollagen is shown in Figure 1-1.

Out of 1050 residues of the collagen α 1(I) chain, the N- and C- terminal end regions are called the telopeptides. These consist of 10-25 residues long sequences each, which are different from the middle portion of the chain. These telopeptides are globular in shape and do not have triple-helical conformation like the body of the molecule. The non-repetitive N- and C-terminal regions do not contain glycine in every third position and are found to be rich in large

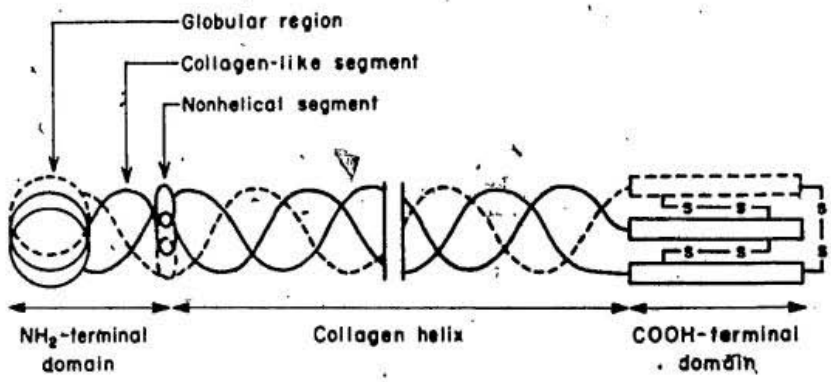


Figure 1-1: The Schematic Representation of Type I Procollagen The molecule is composed of two identical pro-α-1chains (solid lines) and one pro-α-2 chain (dashed line). In addition to the central triple-helical region that gives rise to the collagen molecule, the precursor contains N- and C-terminal nontriple-helical domains. The N-terminal region is composed of a presumably globular region, a short collagen-like segment, and a nontriple-helical region in which cleavages by N-terminal protease occur. Inter-chain disulphide bonds are limited to the C-terminal domain. The short telopeptides at the ends of collagen α-chains represent the residual sequences of the linkage regions between the collagen helix and the terminal domains. (The C-terminal telopeptide is not shown.)

Reproduced from Bornstein and Traub, 1979.

hydrophobic and charged amino acids. Cysteine, which is absent from the triple-helical region of procollagen, is present in both the N-terminal and C-terminal extensions but inter-chain disulphide bonds are limited to the C-terminal domain in type I procollagen (Bornstein and Traub, 1979). Due to the presence of these disulphide bonds in the C-terminal region, this region has been implicated in registering the three procollagen α-chains for triple-helix formation (Bornstein and Traub, 1979; Capaldi and Chapman, 1982). These regions are the sites of lysine residues which are involved in interolecular cross-links (Stoltz *et al.*, 1973; Tanzer, 1976). The telopeptide regions also contain other sequences that are involved in directing the formation of fibrils (Hulmes *et al.*, 1973; Comper and Veis, 1977a, b). Helseth *et al.* (1979), observed that the addition of free amino-terminal telopeptide isolated from the α(I) chain, to a solution of native collagen, specifically enhanced the rate of fibril-formation.

The bulk of the polypeptide chain, about 1011 residues long, consists of the sequences which form the major helical domain of the collagen molecules (Piez, 1976). An examination of these sequences reveal that every third position is occupied by Gly in a very regular fashion throughout the helical domain (Fietzek *et al.*, 1972a, b: Gross, 1976). Consequently, the polypeptide chains of collagen can be considered as repeating triplets of the type (Gly-X-Y). The three positions of (Gly-X-Y) triplet will be denoted as position 1, 2 and 3 respectively, throughout the thesis. The X and Y positions of this repeating triplet can be occupied by a variety of amino acids usually other than Gly. Gly occurs only once in an X position (residue 327 of α1(I)). Proline is found to occupy the X position most frequently, while Hyp occurs only in the Y position (Piez, 1976; Hoffmann *et al.*,

1980). Several of the amino acids are found to have a random distribution in positions 2 and 3 of the (Gly-X-Y) triplet although some residues show preference for a particular position over the other. Some of these unequal distributions are shown to be consistent with intramolecular interactions that could stabilize the molecular structure (Salem and Traub. 1975).

Comparison of sequence data indicates that the α chains from different species or tissues are homologous. Where differences occur, they usually involve conservative substitutions especially in the case of charged amino acids which may be critical in stabilizing the molecular and macromolecular structure. Large hydrophobic residues such as Trp seem to be less critical to the structure (Piez, 1976; Hoffmann *et al.*, 1980).

1.3. Molecular Conformation of Collagen

The molecular structure of collagen constitutes a unique class of supersecondary structures observed in proteins. The now well-known triple-helical conformation of collagen was first proposed by Ramachandran and Kartha (1954, 1955 a,b). This model required the presence of Gly in every first position along the chain. Ramachandran and Kartha's original proposal (1954) of 3 peptide units per turn (n) of the helix was subsequently modified by them to 3.3 units per turn with a unit height (h) of 2.9 A^0 (Ramachandran and Kartha, 1955 a, b). The molecular structure of collagen was deduced from X-ray diffraction data on fibrous collagen and from low-resolution single crystal X-ray diffraction data on the synthetic polypeptide model, namely, $(Pro\text{-}Pro\text{-}Gly)_{10}$ (Okuyama *et al.*, 1972). According to these studies, the collagen molecule consists of 3

individual extended left-handed helices each with a -110° twist and these three helices are coiled in a right-handed superhelix with +30° twist. This structure was earlier called the "coiled-coil" structure of collagen and is analogous to the three strands of a rope wound around each other. Figure 1-2 shows the coiled-coil structure of collagen. With Gly always in the first position and situated at the narrow twist region, such a structure can readily accommodate the rigid and bulky imino residues in the other regions. The presence of Gly in the first position is shown to be important because of its small size and lack of side chain. If the Gly residues are replaced by any other amino acid with a side chain, the chains have to be moved apart and then the inter-chain H-bonds cannot be formed. This results in the destabilization of the triple-helical structure. The number of inter-molecular hydrogen bonds (H-bonds) per tripeptide unit has been controversial (Ramachandran and Ramakrishnan, 1976).

1.3.1. Role of Hydroxyproline in Collagen Conformation

As early as 1955, Gustavson, based on a study of the observed correlation between melting temperature and Hyp content of various collagens, suggested that this imino acid may have a role in stabilizing the collagen structure. Later investigators postulated that the melting temperatures of collagens can be better correlated with the total imino acid content, i.e. Pro and Hyp together (Harrington and Von Hippel, 1961; Josse and Harrington, 1964). However, the actual mode of stabilization was not clear until recently, when Hyp has been shown to play a crucial role in stabilizing the triple-helix of collagen under physiological conditions (Berg and Prockop, 1973 a, c; Sakakibara *et al.*, 1973; Fessler and Fessler, 1974). These studies showed that the transition temperature

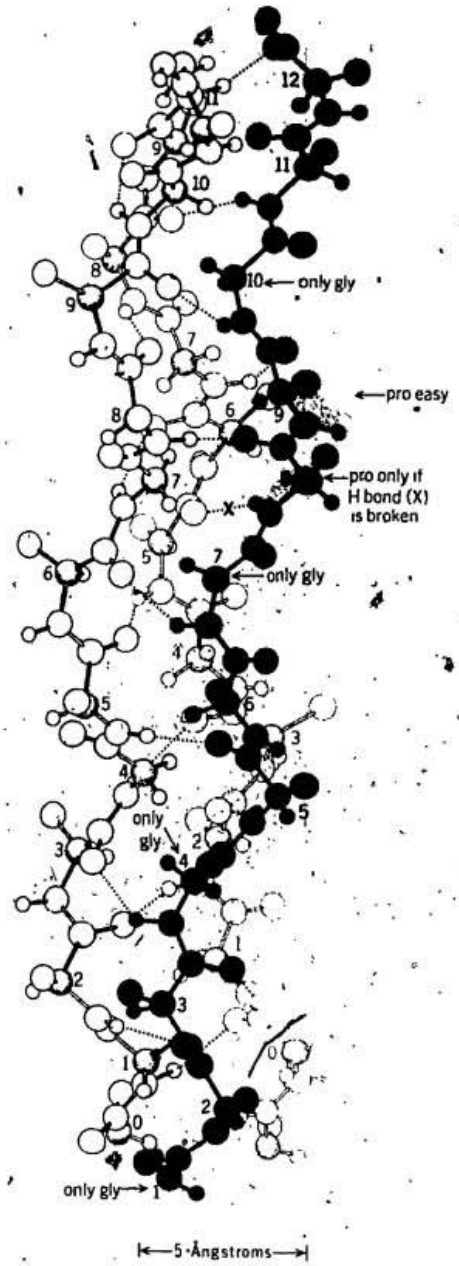


Figure 1-2: The Basic Coiled-coil Structure of Collagen

Three left-handed single-chain helices wrap around one another with a right-handed twist.

Reproduced from Dickerson and Geis, 1969.

(T_m) for the unfolding of the molecule consisting of three non-hydroxylated pro-α-chains is only 24 °C, a value about 15 °C lower than the T_m for the molecules having hydroxylated pro-α-chains. Thus, the unhydroxylated pro-α-chains cannot form triple-helical molecules at the usual physiological temperatures (of the respective species). A certain level of proline hydroxylation seems to be necessary for the formation of the triple-helical collagen molecule that is stable at 37 °C, the body temperature of most mammals.

The role of Hyp in stabilizing the collagen structure has also been investigated by theoretical computations (Bansal *et al.*, 1979). Earlier studies by Ramachandran *et al.* (1973) showed that the γ-hydroxyl group of Hyp cannot form a direct inter-molecular H-bond. But, in a "water-bridged" structure, the γ-hydroxyl group, if it is in a trans-orientation (but not in a cis-orientation), can participate in H-bonding with the bridged water molecule. This molecule also links a carbonyl oxygen in the same chain to an amino group in a neighbouring chain of the triple helix, as shown in the Figure 1-3.

In addition, the γ-hydroxyl group can also form a H-bond with a carbonyl oxygen, in a neighbouring triple-helix in the collagen fibril (Ramachandran *et al*, 1973; Bansal *et al.*, 1979). It is to be noted that neither of these H-bonds can be formed if the Hyp residue occurs in the 2nd position instead of the 3rd in the Gly-X-Y triplet. This model thus may explain the evolutionary selection of Hyp residues in the 3rd position of the Gly-X-Y triplet and also the need for the trans-orientation of γ-OH group at the Cγ atom of proline ring. X-ray and theoretical studies on polyhydroxyproline (Sasisekharan, 1959 b; Bansal *et al.*,

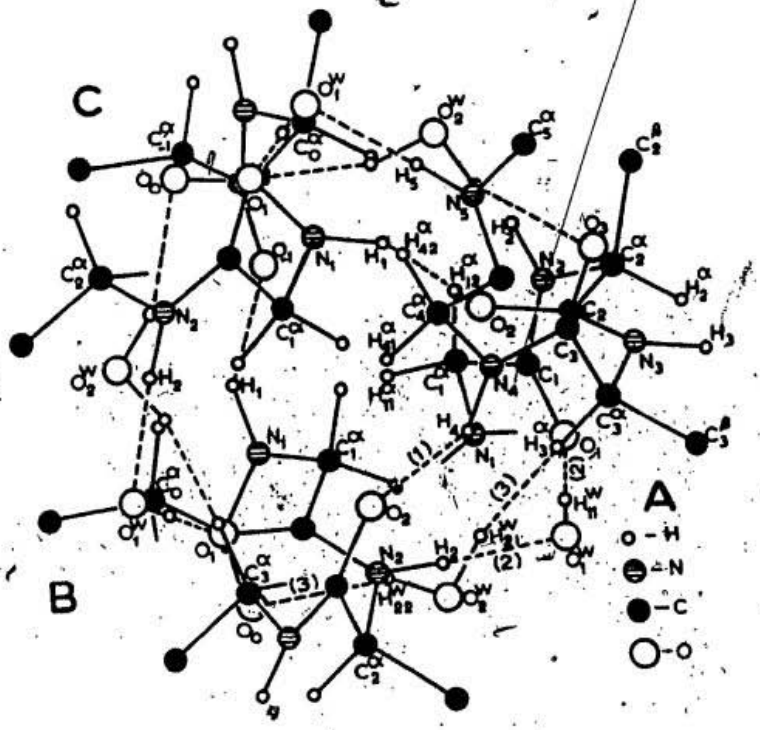


Figure 1- The Water-bridged Structure of Collagen

A projection of the water-bridged structure down the helical-axis. The covalent bonds are shown by solid lines and the hydrogen bonds are shown by dashed lines.

Reproduced from Bansal *et al.*, 1979.

1979; Brahmachari and Ananthanarayanan, 1979) as well as NMR data of Torchia and Lyerla (1974) on $(Hyp\text{-}Gly)_n$ clearly demonstrated that the γ-hydroxyl group of the (*i*)th Hyp residue would be involved in an intra-molecular H-bonding with the carbonyl oxygen of the (*i*-2)th residue in the polypeptide chain. These studies indicate that, in addition to stabilizing the collagen structure due to the stereochemical properties of the rigid pyrrolidine rings, the Hyp residues can also provide additional stability due to H-bonding through their γ-hydroxyl groups.

1.4. Collagen Biosynthesis

1.4.1. Posttranslational Modifications of Collagen Peptide Chains

The posttranslational modification of proteins is now a well-demonstrated phenomenon (Park and Kim, 1975; Freedman and Hawkins, 1980). The modifications often result in the conversion of one form of the protein into another, either covalently or non-covalently. This leads to significant changes in the protein structure with possible changes in its function as well. Examples for non-covalent modifications include association and dissociation of subunits, holozyme formation from prosthetic group and apoenzyme etc. The major types of covalent modifications involve the modifications of amino acid side chains and limited proteolysis. Examples include methylation, glycosylation, phosphorylation, iodination, disulphide-bridge formation, carboxylation, hydroxylation, ADP-ribosylation and thiolation. Besides the structural and/or functional advantage gained by posttranslational modifications, another advantage is the possibility for regulation at different stages, after the transcription and translation of the protein on ribosomes. Thus, a variety of

tissue-specific homologous proteins can be formed that are regulated indepentently of the regulation at the transcriptional or translational level. The main posttranslational events in the biosynthesis of collagen are summarized in Table 1-1.

As shown in Table 1-1, the posttranslational modifications of collagen seem to take place at two locations, intracellular and extracellular. All the intracellular enzymatic processing of the procollagen polypeptides probably occur within the rough endoplasmic reticulum (Prockop *et al.*, 1976). The nascent procollagen polypeptides have signal sequences similar to those in most other "export" proteins (Yamada *et al.*, 1983). These sequences are cleaved during, or shortly after, the translocation across the membrane by signal peptidases.

There are no codons specifying Hyp and Hyl in the genetic code for collagen (Urivetzky *et al.*, 1966) i.e. these amino acids are not incorporated into the collagen polypeptides during the translation of collagen mRNA. However, they are formed by the enzymatic hydroxylation of Pro and Lys residues present in the polypeptide chains after they are translated on the ribosomes. Therefore, the hydroxylation can be considered as a posttranslational modification . However, these reactions may be initiated as a cotranslational event meaning that these reactions are carried out while part of the polypeptide chain is still growing on the ribosomes (Kivirikko and Myllyla, 1980). These reactions are continued within the cisternae of rough endoplasmic reticulum until triple-helix formation of the procollagen chains prevents any further hydroxylation. Proline hydroxylation will be discussd in detail later (Section 1.5). Glycosylation of asparagine (Asn) and

Table 1-1: Posttranslational Events in Collagen Biosynthesis

Process	Enzyme	Significance of the Process
Intracellular:		
1. Removal of the Signal Peptide	Signal Peptidase	Translocation across the membrane
2. 3-hydroxylation of proline.	3-prolyl hydroxylase	Unknown
3. 4-hydroxylation of proline	4-prolyl hydroxylase	Triple-helix formation
4. 5-Hydroxylation of lysine	Lysyl hydroxylase	Glycosylation of Hyl; Stability of cross-links
5. Glycosylation of Hyl	O-Glycosyl transferase	Not known
6. Glycosylation of Propeptides	N-Glycosyl transferase	Not known
7. Chain association	Non-enzymatic	Formation of trimers
8. Disulphide formation	S-S isomerase	Aid in Triple-helix formation
Extracellular :		
9. Removal of N-propeptides	Procollagen N-proteinase	Normal fibrillar morphology
10. Removal of C-propeptides	Procollagen C-proteinase	Fibril formation
11. Ordered aggregation	Non-enzymatic	Fibre formation
12. Cross-link formation	Lysyl oxidase	Stabilization of the fibres

After Kivirikko and Myllyla, 1985.

Hyl may take place in the similar way. It seems that inter-chain disulphide (S-S) bonds are not formed until pro-α-chains are completely synthesized. An important function of the inter-chain S-S bonds between the C-terminal propeptides in procollagens is to direct the association of polypeptide chains and serve as an initiation point for triple-helix formation (Tanzer *et al.*, 1974; Park *et al.*, 1975; Kivirikko and Myllyla, 1985).

The triple-helical procollagen molecules are secreted into extracellular space, the time required for the folding directly affecting that required for secretion. In the extracellular space, their peptide extensions are removed by specific enzymes. The collagen molecules produced by the cleavage of propeptides are shown to have a remarkable tendency for self-assembly and spontaneous formation of fibrils and other ordered structures. Cross-linking of collagens then takes place between different native fibrils. The initial event in the cross-linking is the oxidative deamination of the ϵ-amino group in certain Lys and Hyl residues to the corresponding aldehyde in a reaction catalyzed by lysyloxidase. A single lysyloxidase acts on both Lys and Hyl residues, the activity being greater for Hyl than for Lys (Siegel, 1979). The reactive aldehydes then participate in the formation of various cross-links which stabilize the collagen fibrils to a greater extent.

1.4.2. Role of Conformation in the Posttranslational Processing Reactions of Collagen

The posttranslational processing of procollagen requires at least nine specific enzymatic reactions as shown in the Table 1.1. Almost all the specific processing enzymes demonstrate an unusual relationship to the conformation of the protein being processed. It has been demonstrated that the pro-α-chains must be non-helical in order to serve as substrates for the five intracellular enzymes, the three collagen hydroxylases (prolyl-4-hydroxylase, prolyl-3-hydroxylase and lysylhydroxylase) and the other two glycosyltransferases (Kivirikko and Myllyla, 1985). The triple-helix formation prevents the action of all these enzymes and the folding, in fact, limits the extent of intracellular posttranslational modifications.

Most of the extracellular enzymes also show strict conformational requirements in their substrates. Procollagen N-terminal proteinases act upon the corresponding proteins only if they are in the triple-helical conformation. Lysyloxidase acts on collagen only after it has become aggregated into the native type fibrils (Kivirikko and Myllyla, 1985). From these studies, it seems that the collagen molecule, with its unique conformation, can regulate the activities of the enzymes which can interact with it during its biosynthesis.

The present thesis is concerned with the detailed conformational aspects of proline hydroxylation which will be described in the following sections.

1.5. Proline Hydroxylation in the Biosynthesis of Collagen

The posttranslational hydroxylation of prolyl residues is known to be catalyzed by two separate enzymes :

1. Prolyl-4-hydroxylase (prolyl-glycyl peptide, 2-oxoglutarate:oxygen oxidoreductase, 4-hydroxylating, EC 1.14.11.2) and

2. Prolyl-3-hydroxylase (prolyl-glycyl peptide, 2-oxoglutarate:oxygen oxidoreductase, 3-hydroxylating, EC 1.14.11.4)

In the present thesis, we are concerned with only prolyl-4-hydroxylase ("prolylhydroxylase " is used instead of prolyl-4-hydroxylase hereinafter).

1.5.1. Isolation and Physicochemical Properties of Prolylhydroxylase

Prolylhydroxylase activity has been found in many sources including chick embryo, fetal rat skin (see Cardinale and Udenfriend, 1974 for a review), human fetal skin and placenta (Kuutti *et al.*, 1975). Prolylhydroxylase has also been detected in plant tissues (Sadava and Chrispeels, 1971) and in certain microorganisms (Katz and Li, 1972).

Prolylhydroxylase was first obtained in relatively pure form from chick embryo extract (Halme *et al.*, 1970; Pankalainen *et al.*, 1970) and new born rat skin (Rhoads and Udenfriend, 1970) by conventional procedures involving salt precipitation, ion-exchange and gel filtration chromatography techniques. Subsequently two affinity procedures have been developed. The first one (Berg and Prockop, 1973a) involves affinity binding on a column containing a substrate,

the reduced and methylated collagen, covalently linked to agarose and the elution of the enzyme with a polytripeptide substrate, $(Pro\text{-}Pro\text{-}Gly)_n$. The second procedure is based on the affinity of the enzyme for poly(L-proline) or poly(Pro), a competitive polypeptide inhibitor of the enzyme. The polypeptide is covalently coupled to agarose and the bound enzyme is eluted with the same polypeptide of lower molecular weight followed by gel filtration (Tuderman *et al.*, 1975). This affinity procedure has been recently modified further to include a DEAE-ion exchange chromatography step which efficiently removes the bound poly(Pro) from the enzyme (Kedersha and Berg, 1981).

The relative molecular weight (M_r) of prolylhydroxylase from chicken embryos and from rat and human sources is about 240,000 as found by sedimentation equilibrium and gel filtration studies (Berg and Prockop, 1973a; Tuderman *et al.*, 1975; Kuutti *et al.*, 1975; Risteli *et al.*, 1976). Prolylhydroxylase from these sources has been shown to be a tetramer consisting of two different types of enzymatically inactive monomers α and β subunits, with M_r of about 64,000 and 60,000 respectively (Tuderman *et al.*, 1975; Kuutti *et al.*, 1975; Kedersha and Berg, 1981). Very little is known about the secondary and tertiary structures of prolylhydroxylase. Electron microscopy studies indicated that the monomers are rod-shaped and are joined to form V-shaped dimers which are interlocked to form tetramers, $\alpha_2\beta_2$ (Olsen *et al.*, 1973). Intra-chain S-S bonds seem to be essential for maintaining the native structure and activity (Berg *et al.*, 1979).

Prolylhydroxylase from chicken embryos (Tuderman *et al.*, 1975), rat skin

(Chen-Kiang *et al.*, 1977) and human fetal skin (Kuutti *et al.*, 1975) show very similar amino acid compositions. The enzyme is shown to be highly acidic with an isoelectric point of 4.4 (Pankalainen *et al.*, 1970). The α and β subunits of prolylhydroxylase have been shown to differ in amino acid compositions and carbohydrate contents (Berg *et al.*, 1979; Chen-Kiang *et al.*, 1977).

1.5.2. Cofactors and Cosubstrates of Prolylhydroxylase

Prolylhydroxylase belongs to the class of mixed-function oxygenases. All these enzymes require the same cosubstrates, molecular oxygen, ferrous iron and a reducing agent such as ascorbate (Cardinale and Udenfriend, 1974; Hayashi *et al.*, 1975). The specific requirements for these factors are discussed below.

1.5.2.1. Ferrous Ion

Non-heme iron in the ferrous form is shown to be a very specific requirement for prolylhydroxylase (Hurych and Chvapil, 1965; Prockop and Juva, 1965b; Kivirikko and Prockop, 1967c). However, the requirement for exogenous ferrous iron and the nature of any enzyme-bound iron are controversial (Tuderman *et al.*, 1977a; Nietfield and Kemp, 1980; Hanauski-Abel and Gunzler, 1982). Several divalent cations inhibit the enzyme competitively with respect to Fe^{+2} ions, the most potent one being Zn^{+2} (Rapaka *et al.*, 1976; Tuderman *et al.*, 1977a; Vistica *et al.*, 1977). Many metal chelators such as EDTA and $\alpha\alpha$-dipyridyl inhibit prolylhydroxylase competitively with respect to Fe^{2} ions (Kivirikko and Prockop, 1967c; Juva and Prockop, 1969).

1.5.2.2. α-Ketoglutarate or α-KG

Prolylhydroxylase from all sources is shown to exhibit an absolute requirement for 2-oxoglutarate or α-Ketoglutarate (abbreviated as α-KG hereafter). Oxidative decarboxylation takes place stoichiometrically with the hydroxylation of Pro residues (Rhoads and Udenfriend, 1968). In the absence of the peptide substrate, the enzyme catalyzes an uncoupled decarboxylation of α-KG (Cardinale and Udenfriend, 1974; Tuderman *et al.*, 1977a). This partial reaction requires the same cosubstrates and cofactors as the complete reaction and results in the inactivation of the enzyme. This inactivation is probably due to the oxidation of Fe^{+2} ions (Tuderman *et al.*, 1977a; De Jong and Kemp, 1984; Kivirikko and Myllyla, 1985). Competitive polypeptide substrates are known to increase the rate of uncoupled decarboxylation (Counts *et al.*, 1978), which is about $1/80^{th}$ of the coupled reaction (Tuderman *et al.*, 1977a). Many keto acids and other structural analogues of α-KG are reported to inhibit prolylhydroxylase competitively with respect to α-KG (Tuderman *et al.*, 1977a; Majamaa *et al.*, 1985; Cunliffe and Franklin, 1986).

1.5.2.3. Oxygen

One atom of molecular oxygen is fixed into the hydroxyl group of Hyp (Fujimoto and Tamiya, 1962; Prockop *et al.*, 1963), while the other atom of the oxygen molecule is incorporated into the succinate (Cardinale *et al.*, 1971). The nature of the activated oxygen species has been controversial (see Hanauski-Abel and Gunzler, 1982).

1.5.2.4. Ascorbate

Reports on the specificity of ascorbate requirement for prolylhydroxylase have been quite varied. There is no stoichiometric consumption of ascorbate during the hydroxylation of proline (Tuderman *et al.*, 1977a). However, De Jong and Kemp (1984) have reported that the uncoupled decarboxylation of α-KG is stoichiometrically coupled to the oxidation of ascorbate. The current hypothesis is that the nascent collagen polypeptide chains contain many other non-hydroxylatable sequences and when the enzyme encounters such sequences, it undergoes an uncoupled reaction cycle and becomes inactivated due to the oxidation of ferrous ions (De Jong and Kemp, 1984). The main biological function ascribed to ascorbate in collagen biosynthesis is to "reactivate" prolylhydroxylase after such futile uncoupled reaction cycles, by reducing the enzyme-bound ferric ions back to the ferrous state (Kivirikko and Myllyla, 1985).

1.5.3. Mechanism of Enzymatic Proline Hydroxylation

Several mechanisms have been proposed for the action of prolylhydroxylase (see Hanauski-Abel and Günzler, 1982 for a review). The recently proposed mechanisms are based on the analysis of the kinetic data on substrates and cosubstrates (Tuderman *et al.*, 1977a, b; Myllyla *et al.*, 1977, 1978 and 1984; De Jong and Kemp, 1984). According to these, the enzyme first reacts with ferrous ions forming a complex of E-Fe^{+2} which then activates oxygen, probably to O_2^- ion (or radical) resulting in the reactive E-ferryl-oxo complex. This complex then attacks α-KG leading to its decarboxylation. In the absence of the peptide substrate, the decarboxylation results in the conversion of enzyme-bound ferrous ions to the ferric state. These ferric ions remain bound to the active site, making

the enzyme unavailable for new catalytic cycles until the ferric ions are reduced by the ascorbate.

However, under normal circumstances, when the peptide is present and bound to the enzyme at the active site, the decarboxylation of α-KG leads to the formation of a reactive ferryl ion that acts as the active intermediate in oxygen transfer and hydroxylates the peptide-bound proline or lysine. The coupled decarboxylation takes place at much faster rate when compared to the uncoupled reaction. The reactions are shown schematically in Figure 1-4.

1.6. Substrate Recognition and Specifity of Prolylhydroxylase

Prolylhydroxylase is unusual among other hydroxylases in the same class (except lysylhydroxylase), in that its substrate is a macromolecule. Although the actual chemistry of catalysis may be similar to the other hydroxylases, inherent factors like the sequence and the related local and general conformations of the polypeptide substrate may be expected to play an important role in the substrate recognition and specificity of the reaction. In fact, prolylhydroxylase shows a high degree of substrate specificity both in terms of peptide sequence and conformation. Due to the complexity of its sequence, the analogue of the biological substrate of the enzyme, namely, protocollagen (obtained by preventing the hydroxylation reaction, during collagen biosynthesis in fibroblast cultures) does not permit a clear-cut examination of the regulatory sequences and their conformational features within the substrate molecule. Therefore, synthetic oligo- and polypeptides have been used as substrate models and this has helped to a great extent in understanding the various levels of substrate specificity of prolylhydroxylase.

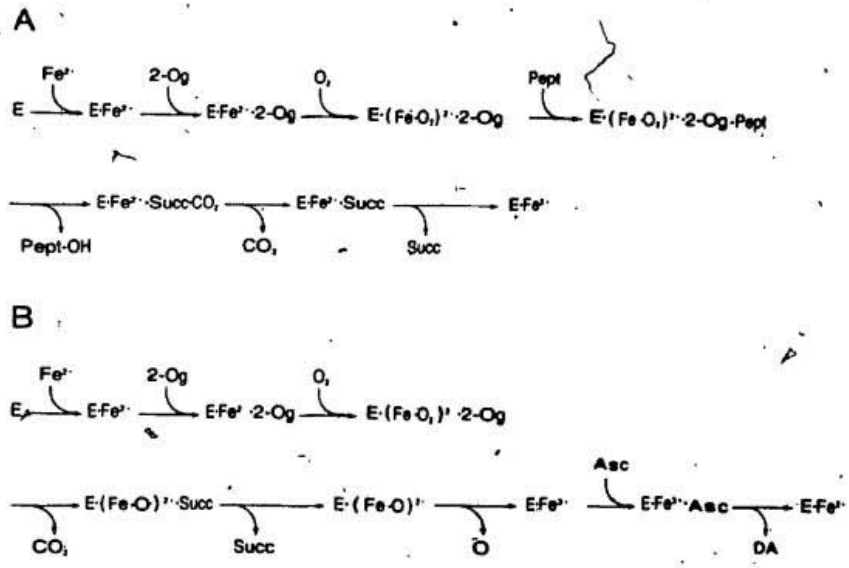


Figure 1-4: Mechanism for Prolylhydroxylase Reaction

Schematic representation of the mechanism for the prolyl-4-hydroxylase reaction. The complete hydroxylation reaction is thought to occur according to scheme (A), in which the order of binding O_2 and the peptide substrate and the order of release of the hydroxylated product and CO_2 are uncertain. In the absence of the peptide, the enzyme catalyzes an uncoupled decarboxylation of 2-oxoglutarate, as shown in scheme (B). E, enzyme; 2-Og, oxoglutarate; Pept-OH, hydroxylated product; Succ, succinate; Asc, ascorbate; Fe^{+2}, ferrous ion; Fe^{+3}, ferric ion; DA, dehydroascorbate.

Adapted from Majamaa *et al.*, 1984.

1.6.1. Specificity for the Primary Structure of the Substrate

The early information about hydroxylatable sequences came from the studies on synthetic tripeptides Ala-Pro-Gly, Pro-Pro-Gly and Pro-Pro-Gly-$NHCH_3$ which were found to be hydroxylated by prolylhydroxylase, while Gly-Pro-Pro, Gly-Pro-Ala or Pro-Gly-Pro were not hydroxylated (Kivirikko *et al.*, 1969; Kikuchi *et al.*, 1969; Suzuki and Koyama, 1969; Kishida *et al.*, 1976). Interestingly, the naturally occuring vasoactive bradykinin, which is totally unrelated to collagen, has been found to be hydroxylated (Rhoads and Udenfriend, 1969). The sequence of this nonapeptide is: Arg-Pro-Pro-Gly-Phe-Ser-Pro-Phe-Arg and it was shown that only the sequence Pro-Pro-Gly is recognized and the proline followed by glycine is specifically hydroxylated by prolylhydroxylase (Rhoads and Udenfriend, 1969).

Synthetic polytripeptides with the structure $(Pro\text{-}Pro\text{-}Gly)_n$ ($n = 5\text{-}20$) were hydroxylated well, the hydroxylation being on the proline residue preceding the glycine (Kivirikko and Prockop, 1967a,c). Similar conclusions were drawn by Hutton *et al.* (1967a,b) based on their studies on protocollagen.

Extensive sequence studies of collagen have indicated that Hyp almost always occurs in the Y position of Gly-X-Y triplet sequence of collagen (Bornstein, 1967a). These studies led to the widely-believed hypothesis that the minimum sequence required for substrate recognition by prolylhydroxylase is the X-Pro-Gly triplet. In agreement with this, polymers with the structure $(Ala\text{-}Pro\text{-}Gly)_n$ where $n > 2$ serve as substrates of the enzyme whereas polymers with the structure $(Pro\text{-}Ala\text{-}Gly)_n$ do not (Kivirikko *et al.*, 1968, 1969).

Thus, it appears that although the repeating triplet sequences of collagen are considered as Gly-X-Y, prolylhydroxylase seems to "read" the peptide sequence in the order -X-Pro-Gly- (Prockop *et al.*, 1976). In accordance with this, Hyp residues are almost always found to be followed by Gly residues. However, more recently, a few exceptions to this have been reported. In the human C1q, the sequence X-Hyp-Ala was noted twice indicating that the specificity of prolylhydroxylase for X-Pro-Gly sequence may not be absolute and that Gly, in some cases at least, may be replaced by Ala (Reid, 1977; Reid and Thompson, 1978). Bhatnagar and Rapaka (1976) also reported that Gly can be replaced in synthetic peptide substrates by β-Ala, an amino acid that contains an additional methylene group in the backbone but does not contain a side chain.

1.6.1.1. Effect of Adjoining Residues on Proline Hydroxylation

It appears that the adjoining residue X on the N-terminal side of the hydroxylatable proline residue affects the hydroxylation reaction in the X-Pro-Gly sequence in collagen and related polypeptide substrates of prolylhydroxylase. Studies with polytripeptides of the structure $(X\text{-}Pro\text{-}Gly)_n$ where X is Ala, Pro, Val, Leu or Gly have indicated that Pro or Ala in the X position make them good substrates, Pro being better than Ala (Kivirikko and Prockop, 1967a,c; Prockop *et al.*, 1967; Kivirikko *et al.*, 1968). Leu in the X position seems to reduce the extent of hydroxylation which is further decreased when the X position is occupied by Val (Kivirikko *et al.*, 1972; Rapaka *et al.*, 1978). Similarly, the presence of sarcosine or Sar (N-methyl glycine) or Gly in the X position is shown to completely prevent hydroxylation (Kivirikko *et al.*, 1969; Rao and Adams, 1978; Rapaka *et al.*, 1978).

Other investigations with polypeptides of the $(X\text{-}Pro\text{-}Gly)_n$ type where X is Arg or Glu, have indicated that the positively charged Arg is better than negatively charged Glu, although both of them are unfavourable when compared to Pro (Prockop *et al.*, 1976). It appears that the main effect of the X residue is on the maximal velocity rather than on the binding of the substrate to the enzyme (Kivirikko and Myllyla, 1980). It is known that the side chain of the X residue preceding Pro can interact with the δ-methylene of the pyrrolidine ring (Schimmel and Flory, 1968). Such interactions may influence the stereochemistry of the X-pro peptide bond and subsequently govern the polytripeptide-enzyme interactions (see section 1.8.1)

Data on the effect of residue Y (C-terminal to Gly) in the X-Pro-Gly-Y tetrapeptide sequence on proline hydroxylation are not available, although one might expect such an influence.

1.6.1.2. Effect of Medium-range Side Chain Interactions on Proline Hydroxylation

Earlier studies with bradykinin and its analogues revealed the effects of side chains of residues farther removed from the hydroxylatable proline. Bradykinin contains nine amino acids with a single Pro-Pro-Gly triplet and it has about the same K_m value (1500 μM) as $(Pro\text{-}Pro\text{-}Gly)_5$ when expressed in the molar concentration of the (Pro-Pro-Gly) unit, as shown in Table 1-2. This affinity of prolylhydroxylase towards bradykinin has been attributed to the presence of the two arginyl residues on either end of the molecule. Addition of another Arg to the N-terminal end of bradykinin was found to reduce the K_m to 750 μM, whereas

the addition of Gly to the same position caused the K_m to increase slightly (Table 1-2). Both modifications increased V_{max}. In contrast, addition of Glu to the N-terminal end was found to reduce the value of both K_m and apparent V_{max} (McGee *et al.*, 1971). Similar studies on the interaction of several other bradykinin analogues with prolylhydroxylase have demonstrated an effect on proline hydroxylation, of amino acid residues beyond the triplet to be hydroxylated (McGee *et al.*, 1971).

Kivirikko *et al.* (1972) have compared Arg-Gly-$(Pro\text{-}Pro\text{-}Gly)_5$ and Glu-Gly-$(Pro\text{-}Pro\text{-}Gly)_5$ with $(Pro\text{-}Pro\text{-}Gly)_5$ as substrates of prolylhydroxylase. As shown in Table 1-2, the K_m value for the Arg containing peptide was only about one-half of that for $(Pro\text{-}Pro\text{-}Gly)_5$, but the extent of hydroxylation was not affected. On the other hand, the presence of Glu in polypeptides appears to have little effect on their binding to the enzyme, although the extent of hydroxylation was significantly reduced, when compared to $(Pro\text{-}Pro\text{-}Gly)_5$. Similarly, comparison of $(Arg\text{-}Pro\text{-}Gly)_n$ and $(Glu\text{-}Pro\text{-}Gly)_n$ with $(Pro\text{-}Pro\text{-}Gly)_n$ (as substrates of prolylhydroxylase) indicates that the K_m value for the Arg containing peptide was decreased to 80 μM from 350 μM obtained in the case of $(Pro\text{-}Pro\text{-}Gly)_n$. On the other hand, the K_m value for $(Glu\text{-}Pro\text{-}Gly)_n$ was decreased only to 280 μM. These studies indicate that the enzyme-polypeptide substrate interaction is influenced by amino acids present in other parts of the peptide chain.

The observed effect of increased binding due to a positively charged residue in the vicinity seems to be meaningful in the light of the highly anionic nature of

Table 1-2: Effect of Medium-Range Interactions on Proline Hydroxylation

**Substrate	K_m μM	V_{max} nmoles/hr	% Hydroxy -lation	Ref.
$(P\text{-}P\text{-}G)_n$*	1500	n.a	n.a	1
Bradykinin*	1500	18.0	n.a	1
Arginyl-bradykinin*	750	136.8	n.a	1
Glycyl-bradykinin*	1900	97.2	n.a	1
Glutamyl-bradykinin*	80	5.0	n.a	1
$(P\text{-}P\text{-}G)_n$	350	n.a	100	2
$R\text{-}G\text{-}(P\text{-}P\text{-}G)_n$	170	n.a	115	2
$E\text{-}G\text{-}(P\text{-}P\text{-}G)_n$	400	n.a	75	2
$(R\text{-}P\text{-}G)_n$	80	n.a	-	3
$(E\text{-}P\text{-}G)_n$	280	n.a	-	3

** One letter symbol used for amino acids in synthetic polypeptides

* K_m values expressed in μM concentration of the X-Pro-Gly tripeptide units

n.a: Not available

' K_m values expressed in μM concentration of the peptide

1, McGee *et al.*, 1971

2, Kivirikko *et al.*, 1972

3, Prockop *et al.*, 1976

the enzyme (pI 4.4) at physiological pH. However, by the same token, Glu should have an opposite effect which is seen in the case of Glu-Gly-(Pro-Pro-Gly)$_n$ but not with bradykinin analogue.

1.6.2. Effect of Substrate Chainlength on Proline Hydroxylation

Enzymes which act upon macromolecular substrates are usually different from those operating on smaller substrates. The former, in general, are known to contain a relatively larger active site with subsites for the binding of monomeric units of the substrate, the catalytic site being located at a unique position within the active site (Berger and Schechter, 1970). Enzymes like DNA polymerase I (McClure and Jovin, 1975; Sherman and Gefter, 1976) and nucleic acid methylases (Kerr and Borek, 1973) are known to exhibit multiple equivalent sites of binding. Similarly, prolylhydroxylase with its long nascent procollagen polypeptide substrate of about 1000 amino acids might be expected to exhibit binding subsites and multi-point attachments. Some insights into the enzyme's complex behaviour has been obtained from the studies discussed below.

Studies on protocollagen, with a M_r of 150,000 have demonstrated that it is bound by prolylhydroxylase with the highest affinity among all the substrates which have been ever studied (Kivirikko and Prockop, 1967c; Berg and Prockop, 1973b). It has a K_m of about 2 nM and a K_{cat} value equal to 4-6 sec^{-1} (Berg and Prockop, 1973c). The immediate explanation for this high affinity was in terms of chainlength, although part of the effect may be due to the varied and complex amino acid composition of the substrate. Thus, protocollagen's affinity towards the enzyme may be due to its capability to bind completely and efficiently the

enzyme's large active site by way of multi-point attachment at the different binding subsites.

Studies with synthetic polypeptides with different molecular weights have indicated a marked effect of chainlength, such that the higher molecular weight polymers exhibited higher affinity for prolylhydroxylase (as expressed by lower K_m values), although all of them underwent hydroxylation with similar maximal velocities (Prockop and Kivirikko, 1969). The higher affinity of the enzyme for larger peptides may be due to better binding at several binding subsites on the enzyme, as in the case of protocollagen.

Interesting experiments by Kivirikko *et al.* (1971) and Berg *et al.* (1977) have demonstrated that the different tripeptide units in $(Pro\text{-}Pro\text{-}Gly)_5$ and $(Pro\text{-}Pro\text{-}Gly)_{10}$ are hydroxylated to different extents and the penultimate triplet from the N-terminus is the most hydroxylated. This preferential hydroxylation has been explained in terms of an asymmetric active site in which binding subsites are located adjacent to but not symmetrical with the catalytic site (Berg *et al.*, 1977). Quite recently, de Waal *et al.* (1985) demonstrated the presence of possible binding subsites in prolylhydroxylase. They first covalently blocked the active site with a photoaffinity label, N-(4-azido-2-nitrophenyl)-glycyl-$(Pro\text{-}Pro\text{-}Gly)_5$ and then checked whether the enzyme still bound to a polyproline (PP-II) affinity column (section 1.5.1). Interestingly, the covalently bound photoaffinity label did not impair the binding of the enzyme to the PP-II column, although it inhibited the hydroxylation of synthetic peptides suggesting that there are possible binding subsites which may be located adjacent to but not contiguous with the catalytic site.

1.6.3. Role of Substrate Conformation in Proline Hydroxylation

One of the interesting observations about prolylhydroxylase action has been the marked effect of the conformation of the peptide substrate. This effect can be at two levels:

1. Overall conformation of the substrate which may affect the binding and

2. Conformation of the minimal sequence i.e. -X-Pro-Gly- required at the catalytic site.

Initial evidence for the former has been obtained from studies on the hydroxylation of the Hyp-deficient cuticle collagen (cuticlin) of *Ascaris* worms. This collagen could be further hydroxylated only if it were thermally denatured prior to hydroxylation (Fujimoto and Prockop, 1968; Rhoads and Udenfriend, 1968). Sequence studies on various collagens indicated that many prolyl residues in the Y position of the Gly-X-Y triplets remained unhydroxylated (Bornstein, 1967a, b; Fietzek *et al.*, 1972a,b). However, many of these residues could be further hydroxylated by prolylhydroxylase, if the collagens were thermally denatured first before the interaction with the enzyme (Rhoads *et al.*, 1971). The explanation offered for these observations was that the enzyme cannot hydroxylate peptide substrates if these are in the triple-helical conformation. Some earlier studies with the synthetic peptides also suggested the inhibitory effect of their triple-helical conformation on hydroxylation (Kikuchi *et al.*, 1969; Kivirikko *et al.*, 1972).

These data on the unsuitability of the triple-helical collagen molecule as a

substrate for prolylhydroxylase were reconciled with the data obtained *in vivo* on proline hydroxylation by Prockop and his colleagues (1976). These investigators showed that the T_m (melting temperature) of protocollagen was 24 °C and thus the protein that was thought to be in the triple-helical conformation during the posttranslational hydroxylation at the body temperature *viz.* 37 °C was actually in the denatured (random coil) conformation. This would then explain its ability to serve as a good substrate *in vivo* (Berg and Prockop, 1973 b,c; Jimenez *et al.*, 1973). Once this was understood, it was possible to obtain protocollagen at a temperature below its melting temperature mainly in the triple-helical conformation; in this form, it did not act as a substrate for prolyhydroxylase under the conditions in which non-helical protocollagen was rapidly and completely hydroxylated (Berg and Prockop, 1973 b,c; Murphy and Rosenbloom, 1973).

The above studies have shed light on the gross conformational requirement for the proline hydroxylation in peptide substrates. However, to understand fully the conformational aspects of prolylhydroxylase substrate specificity, factors governing the formation of triple-helical structure and the basis for the differential recognition of helical and non-helical conformations in the substrates have to be made clear. More importantly, the localized conformations, if any, at the actual hydroxylated sequences are to be taken into consideration in order to define the specificity in more certain terms. Together, these studies should provide valuable information, in conformational terms, regarding the factors that regulate proline hydroxylation in collagen. Synthetic polypeptide models and small oligopeptides have been found extremely useful in such studies, as described in the next two sections.

1.7. Polypeptide Substrates of Prolylhydroxylase

The natural substrates of prolylhydroxylase, protocollagen polypeptide chains are polymers of glycine-led tripeptide units in which the second and/or third positions are usually occupied by imino acids. For recognition and complex formation with the polypeptide-binding site of prolylhydroxylase, the model compounds must exhibit some characteristic features of the natural substrate. Therefore, a number of homo- and heteropolypeptide substrates of $(Gly\text{-}X\text{-}Pro)_n$ and $(Gly\text{-}Pro\text{-}X)_n$ type have been synthesized where X is Ala, Ser, Sar or Leu. These synthetic peptides have been studied for their conformational features as well as their capability to interact with prolylhydroxylase, either by way of hydroxylation or by competitive inhibition with respect to the natural substrate (Doyle *et al.*, 1971; Brown *et al.*, 1972; Bhatnagar and Rapaka, 1976). As a background for understanding the data on proline hydroxylation of synthetic polypeptide models of collagen, an outline of their structural features is presented below.

1.7.1. The Structural Features of Polypeptide Substrates and Inhibitors

Because of the large imino acid content of collagen, the stereochemical interactions of the imino peptide bonds dominate the secondary structure of collagen. Synthetic polypeptides containing significant proportions of these residues are therefore expected, in general, to exhibit many conformational characteristics of the collagen molecule.

1. Poly(Pro): This homopolymer exists in two conformations: poly(Pro) I (PP I) occurs in a right-handed helix with all the peptide bonds in

cis-configuration and poly(Pro) II (PP-II) takes up a left-handed helix with all the peptide bonds in trans-configuration (Sasisekharan, 1959a). The latter conformation has been found to be very similar to the single chain conformation of collagen. As will be described later, only PP-II polymers interact with prolylhydroxylase (Kivirikko *et al.*, 1967).

2. Poly(4-hydroxy-L-Proline) or Poly(Hyp): This homoploymer shares many structural features with poly(Pro) although some basic differences exist. Poly(Hyp) also takes up a left-handed helical structure with all the peptide bonds in trans-configuration. In this structure, the γ-hydroxyl group of an (i)th Hyp residue is directly H-bonded to a carbonyl oxygen of the (i-2)th residue in the same polypeptide chain (Bansal *et al.*, 1979). The poly(Hyp) structure is considerably more stable than the PP-II helix in aqueous solution (Mattice and Mandelkern, 1970) probably due to this intra-molecular H-bonding.

3. Poly(Gly-Pro-X): In the polytripeptides of the $(Gly\text{-}Pro\text{-}X)_n$ type, where X is Ala, Ser, Leu or Sar, it has been observed that the presence of Pro in the second position (i.e. C-terminal to Gly) favours the formation of the triple-helical conformation (Brown *et al.*, 1972; Scatturin *et al.*, 1975; Ananthanarayanan *et al.*, 1976).

4. Poly(Gly-X-Pro): In these types of polytripeptides, where X is again Ala, Ser, Leu or Sar, it has been observed that the presence of Pro in the third position (i.e. N-terminal to Gly) results in the unordered structure especially in solution (Doyle *et al.*, 1971; Scatturin *et al.*, 1975; Ananthanarayanan *et al.*, 1976).

5. Poly(Gly-Pro-Pro): This polytripeptide has been extensively studied (Engel, 1967; Ramachandran *et al.*, 1968; Yonath and Traub, 1969; Kobayashi *et al.*, 1970, 1977; Sakakibara *et al.*, 1972; Okuyama *et al.*, 1976) and found to exist in triple-helical conformation. It is one of the best models for collagen. Oligomers (penta- and decamers) of $(Gly\text{-}Pro\text{-}Pro)_n$ have been shown to have less stability compared to (Gly-Pro-Hyp) of same chain length, though both of them adopt the triple-helical conformation in solution (Sakakibara *et al.*, 1968, 1973; Sutoh and Noda, 1974a). This extra stability of $(Gly\text{-}Pro\text{-}Hyp)_n$ is considered to be strong evidence for the hypothesis that Hyp residues stabilize the triple-helical conformation by forming H-bonds through their γ-hydroxyl groups (Ramachandran *et al.*, 1973; Bansal *et al.*, 1979; Engel *et al.*, 1977).

1.7.2. Proline Hydroxylation in Polypeptide Models of Collagen

As described above, the $(Gly\text{-}X\text{-}Pro)_n$ polytripeptides usually occur in random or disordered conformation in aqueous medium, while $(Gly\text{-}Pro\text{-}X)_n$ polypeptides exist in rigid, extended conformation. In general, the former type of peptides have been shown to be substrates undergoing hydroxylation and/or inhibitors of prolylhydroxylase. On the other hand, polypeptides like the $(Gly\text{-}Pro\text{-}Pro)_n$ and the $(Gly\text{-}Pro\text{-}Ala)_n$, which exhibit highly ordered collagen-like conformation in solution (Engel *et al.*, 1966; Brown *et al.*, 1972), were found to behave differently towards prolylhydroxylase. The $(Gly\text{-}Pro\text{-}Pro)_n$ exhibits strong interaction with the enzyme acting, both as a substrate and a competitive inhibitor (Kivirikko and Prockop, 1967a, b; Hutton *et al.*, 1968), whereas the $(Gly\text{-}Pro\text{-}Ala)_n$ showed very little complex formation (Kivirikko *et al.*, 1969). Similarly, the $(Gly\text{-}Pro\text{-}Sar)_n$ which exhibits collagen-like conformation in solution

(Ananthanarayanan *et al.*, 1976), did not interact with prolylhydroxylase (Bhatnagar and Rapaka, 1976). The homopolymer PP-II, which can be considered as $(Pro\text{-}Pro\text{-}Pro)_n$, interacted very strongly with prolylhydroxylase. It bound with very high affinity to the enzyme so that it could be used as a ligand, as part of affinity column material in the purification of prolylhydroxylase. It was not hydroxylated, although it acted as a potent competitive inhibitor (Kivirikko and Prockop, 1967 b,c; Prockop and Kivirikko, 1969). In contrast, PP-I did not interact with the enzyme (Kivirikko *et al.*, 1967). Poly(Hyp) and poly-O-acetyl-(Hyp), which share many structural features with poly(Pro) II (Sasisekharan, 1959a, b; Bansal *et al.*, 1979) did not show any significant interaction (Bhatnagar and Rapaka, 1976) with prolylhydroxylase. It was later pointed out that small conformational differences exist between the Pro and Hyp homopolymers (Torchia *et al.*, 1972; Torchia and Lyerla, 1974) and such differences may be responsible for the absence of interaction of the Hyp polymers with the enzyme (Bhatnagar and Rapaka, 1976). Since the OH-group of poly-O-acetyl(Hyp) is blocked, any interaction between the OH-group of Hyp and the backbone is absent. The inability of both poly(Hyp) and poly-O-acetyl(Hyp) to bind to prolylhydroxylase thus indicates that the substitution at the γ-carbon atom, rather than the backbone conformation, abolishes the interaction. These studies may offer, at least in part, an explanation for the reduced affinity of partially hydroxylated procollagen towards prolylhydroxylase. The $(Gly\text{-}Pro\text{-}Pro)_n$ polymers acted as good substrates only when they were thermally denatured prior to the interaction with prolylhydroxylase i.e. in their unordered forms. Table 1-3 summarizes the substrate specificity of prolylhydroxylase

Table 1-3: Substrate Specificity of Prolylhydroxylase

Substrate	Binding*	Hydroxylation
Native Collagen	Yes	No
Denatured Collagen	Yes	Yes
Free Proline	No	No
Gly-Pro-Pro	No	No
Pro-Pro-Gly-$NHCH_3$	Yes	Yes
Polyproline I	No	No
Polyproline II	Yes	No
Poly(Hyp)	No	No
Poly(O-Ac-Hyp)	No	No
$(Gly\text{-}Pro\text{-}Pro)_{1\text{-}5,\ 10,\ 20}$	Yes	Yes
$(Gly\text{-}Pro)_n$	Yes	No
$(Gly\text{-}Gly\text{-}Pro)_n$	Yes	No
$(\beta\text{-}Ala\text{-}Pro\text{-}Pro)_n$	Yes	Yes
$(Gly\text{-}Pro\text{-}Ala)_n$	No	No
$(Gly\text{-}Ala\text{-}Pro)_n$	Yes	Yes
$(Gly\text{-}Pro\text{-}Sar)_n$	No	No
$(Gly\text{-}Sar\text{-}Pro)_n$	Yes	No
$(Gly\text{-}Pro\text{-}Leu)_n$	No	No
$(Gly\text{-}Leu\text{-}Pro)_n$	Yes	Yes
$(Gly\text{-}Pro\text{-}Gln)_n$	No	No
$(Gly\text{-}[illegible]l\text{-}Pro)_n$	Yes	Yes
$(Val\text{-}Pro\text{-}Gly\text{-}Val\text{-}Gly)_n$	Yes	Yes
$(Val\text{-}Ala\text{-}Pro\text{-}Gly\text{-}Val\text{-}Gly)_n$	Yes	Yes
$(Ala\text{-}Pro\text{-}Gly\text{-}Gly)_n$	Yes	Yes
$(Val\text{-}Pro\text{-}Gly\text{-}Gly)_n$	Yes(?)	No
$(Val\text{-}Ala\text{-}Pro\text{-}Gly)_n$	Yes	Yes
Bradykinin	Yes	Yes

* as determined by the ability to inhibit the enzyme.

From Ananthanarayanan, 1983a.

towards polytripeptide models. Native collagen and other substrates are included for the ease of comparison.

The earlier expectation was that understanding the conformational features of the above polymers (Table 1.3) on one hand and their interaction with prolylhydroxylase on the other, would shed light on the conformational requirements for proline hydroxylation. In particular, since $(Gly\text{-}Pro\text{-}X)_n$ type polymers are not hydroxylated but $(Gly\text{-}X\text{-}Pro)_n$ type polymers are, knowing their conformational characters was expected to solve the problem. The polypeptide data however, did not provide any more insights into the conformational criteria for proline hydroxylation than what were already obtained from the studies on native and denatured collagen or on $(Pro\text{-}Pro\text{-}Gly)_n$ in triple-helical and unordered forms. That is, these studies only reiterated the conclusion that a "flexible" rather than "rigid" conformation is needed for enzymatic proline hydroxylation. At this stage, the problem still remained to be solved and resort was made to the structural analysis of simpler peptide models. Later studies on di- and tripeptides provided important insights into the basis for the intrinsic conformational differences between $(Gly\text{-}Pro\text{-}X)_n$ and $(Gly\text{-}X\text{-}Pro)_n$, by underscoring the importance of the conformational features of the fundamental tripeptide unit itself (see section 1.8.2).

1.7.3. Conformational Criteria for Proline Hydroxylation

Since prolylhydroxylase seems to recognize the (X-Pro-Gly) segments in the substrates, and the (X-Pro-Gly) polymers are shown to be predominantly in non-helical conformation, it is important to study the conformational aspects of the (X-Pro-Gly) sequence in terms of contributions from the X-Pro and Pro-Gly segments.

1.7.4. Conformation of the X-Pro Segment

Theoretical studies on the X-Pro segments have been carried out by Schimmel and Flory (1968) and Rapaka *et al.* (1978). In a polypeptide sequence of the $(X\text{-}Pro\text{-}Gly)_n$ type, the conformational range of the X-Pro peptide unit is restricted by steric interactions between the side chain of the X residue and the proline ring atoms (Schimmel and Flory, 1968). These interactions can influence the conformation of the peptide and hence its interaction with the enzyme. Studies with the polytripeptide models of the $(X\text{-}Pro\text{-}Gly)_n$ type, (where X is Gly, Ala, Leu, Val, Phe, Sar or Pro) have indicated that the nature of the X residue exerts an influence on the extent of the enzyme-peptide interaction (Rapaka *et al.*, 1978). The interaction in terms of the hydroxylation of these polypeptides by prolylhydroxylase is shown in Table 1-4.

As seen from this table, hydroxylation takes place in the decreasing order: Pro > Ala > Leu > Val. On the other hand, Gly, Phe and Sar when present in the X position impaired the hydroxylation. These results have been interpreted by Rapaka *et al.* (1978) in terms of the conformational features of the X-Pro segment in the above polytripeptides, as obtained from conformational energy

Table 1.4: Hydroxylation of Polypeptide Models of Collagen

Polymer	Percentage of Susceptible Pro hydroxylated
$(Pro\text{-}Pro\text{-}Gly)_n$	30.0
$(Gly\text{-}Pro\text{-}Gly)_n$	<2.0
$(Ala\text{-}Pro\text{-}Gly)_n$	16.0
$(Leu\text{-}Pro\text{-}Gly)_n$	10.2
$(Val\text{-}Pro\text{-}Gly)_n$+	5.0
$(Phe\text{-}Pro\text{-}Gly)_n$*	<3.0
$(Sar\text{-}Pro\text{-}Gly)_n$	<2.0
$(Pro\text{-}Pro\text{-}Pro)_n$	<2.0

All peptides had a molecular weight of about 4,000 on the basis of gel filtration.

+ $(Val\text{-}Pro\text{-}Gly)_n$ was moderately soluble in water

* $(Phe\text{-}Pro\text{-}Gly)_n$ was insoluble in water

After Rapaka *et al.*, 1978.

calculations. In the case of those peptides that were hydroxylated, the energetically favoured conformations predominate within a small area of the ψ_1-ψ_2 plot, where ψ_1 [-C^{α}_{x}--C(=O)-] for the residue X is in the range of 100 ± 40^{o} and ψ_2 [-C^{α}_{pro}--C(=O)-] for Pro is 130 ± 30^{o}. Where energetically favourable conformations lie outside this range of ψ_1 and ψ_2, hydroxylation does not occur. It was also proposed that, for proline hydroxylation, the above proposed critical range of conformations is necessary and the extent of reaction may depend on this range of conformations.

Based on the total range of conformations, Rapaka *et al.* (1978) found $(Pro\text{-}Pro\text{-}Gly)_n$, $(Ala\text{-}Pro\text{-}Gly)_n$ and $(Leu\text{-}Pro\text{-}Gly)_n$ to have their energy minima (for the X-Pro unit) well within the range of permissible conformational area. For $(Val\text{-}Pro\text{-}Gly)_n$ and $(Phe\text{-}Pro\text{-}Gly)_n$, although the energetically favourable conformations lie within the permissible conformational area, the decreased hydroxylation observed may be due to the insolubility of these peptides in aqueous medium, under the reaction conditions. For $(Sar\text{-}Pro\text{-}Gly)_n$ and poly(Pro) (viewed as $(Pro\text{-}Pro\text{-}Pro)_n$), the conformations favoured lie outside the permissible area and hence they are unable to undergo hydroxylation (Rapaka *et al.*, 1978).

1.7.5. Conformation of the Pro-Gly Segment: Earlier Data

Turning now to the Pro-Gly unit, the near-absolute requirement for Gly in the hydroxylatable sequence X-Pro-Gly, both in collagen and in synthetic substrates, suggests that the hydroxylation of Pro residues may involve some highly specific stereochemical properties of this unit. In the repeating sequence of

$(Gly\text{-}X\text{-}Pro)_n$, Gly following the Pro residue contributes to the conformational freedom because of its lack of side chain (Schimmel and Flory, 1968). Studies on $(\beta\text{-}Ala\text{-}Pro\text{-}Pro)_n$ by Bhatnagar and Rapaka (1975) showed that this compound exhibits conformational characteristics very similar to that of $(Gly\text{-}Pro\text{-}Pro)_n$ and both undergo hydroxylation with comparable rates and affinity. This suggests that Gly, at least in some cases, can be replaced by β-Ala because of the large freedom of rotation retained by this residue. Studies on $(Gly\text{-}Gly\text{-}Pro)_n$, $(Gly\text{-}Pro\text{-}Sar)_n$ and $(Gly\text{-}Sar\text{-}Pro)_n$ have further highlighted the importance of conformational freedom at the Pro-Gly bond (Ananthanarayanan *et al.*, 1976), by comparing it with the Pro-Sar bond. These studies have demonstrated that when Sar occurs on the C-terminal side of Pro, it affects the conformational properties of the peptide to a considerable extent in comparison to the case in which it occurs on the N-terminal side of the Pro residues. The change in the conformational features is also reflected in its altered interaction with prolylhydroxylase (Bhatnagar and Rapaka, 1976), suggesting the important role of the Pro-Gly sequence and its conformational nature in proline hydroxylation.

The above studies indicated the role of the X residue on the hydroxylation of Pro in the X-Pro-Gly sequence in conformational terms and that of Gly in contributing to the conformational freedom of Pro-Gly segments. However, the basis for the selective hydroxylation of the "disordered" conformation of the natural and the synthetic polypeptides is not understood in conformational terms, from the above described studies (although the X-Pro-Gly sequences are present in both triple-helical as well as disordered conformation of the peptide substrates).

1.7.6. Further Studies on the Conformation of the X-Pro and Pro-Gly Segments

With a view to understand the conformational basis for the specificity of prolylhydroxylase towards the X-Pro-Gly sequences in the disordered state of the polypeptide substrates, Brahmachari and Ananthanarayanan (1979) examined the conformation of the X-Pro and Pro-Gly segments in peptides and proteins, on the basis of the then available data. These authors noted an important basic difference between the conformations of the X-Pro and Pro-Gly segments. Thus, from theoretical conformational energy analysis (Stimson *et al.*, 1977; Zimmerman and Scheraga, 1977), it has been shown that, a Gly-Pro sequence energetically favours an open, extended conformation with $\phi_{Gly} = 178^{o}$, $\psi_{Gly} = 175^{o}$; $\phi_{Pro} = -75^{o}$ and $\psi_{Pro} = 79^{o}$, and can readily adopt the extended PP-II conformation (Sasisekharan, 1959a). On the other hand, a Pro-Gly sequence favours a folded or bent conformation, similar to the β-turn (Venkatachalam, 1968), with $\phi_{Pro} = -75^{o}$, $\psi_{Pro} = 110^{o}$, $\phi_{Gly} = 80^{o}$ and $\psi_{Gly} = 10^{o}$. Figure 1-5 shows the conformations of Pro-Gly and Gly-Pro sequences.

This led Brahmachari and Ananthanarayanan (1979) to provide an explanation for the intrinsic conformational differences between $(Gly\text{-}X\text{-}Pro)_n$ and $(Gly\text{-}Pro\text{-}X)_n$ polytripeptides in terms of the Pro-Gly or Gly-Pro repeating sequence present in the polymer. The $(Gly\text{-}X\text{-}Pro)_n$ with (-Pro-Gly-) segments would tend to take up a folded conformation, whereas the $(Gly\text{-}Pro\text{-}X)_n$ due to the presence of (-Gly-Pro-) segments tend to exist in an extended (similar to PP-II) conformation. Thus, the latter can readily form the collagen-like triple-helical

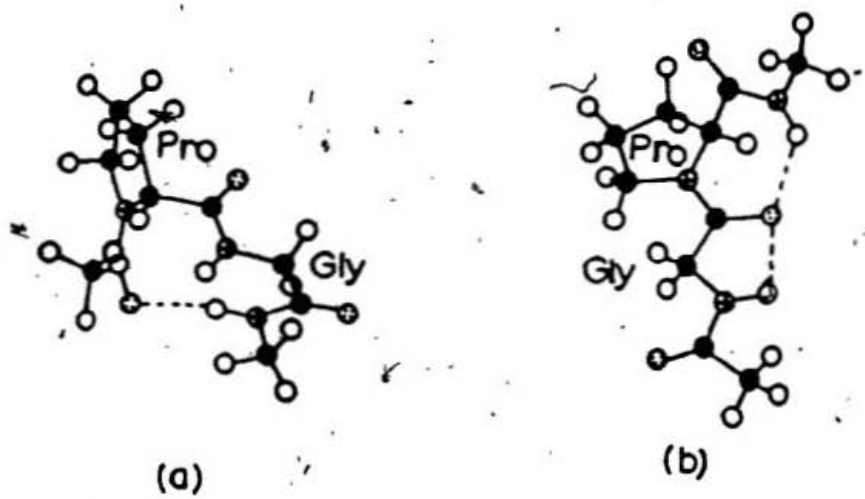


Figure 1-5: Minimum Energy Conformations of Ac-Pro-Gly-$NHCH_3$ (Left) and Ac-Gly-Pro-$NHCH_3$ (Right) with all trans peptide unit

(From Stimson *et al.*, 1977)

structure, while the former would behave like a flexible (though not fully random) molecule.

In order to further understand the conformational characteristics of the Pro-Gly segment in peptides, particularly with respect to the role of the X residue at the C-terminal side of Gly, a detailed study of the conformation of N-Acetyl-Pro-Gly-X-OH where X is Gly, Ala, Ile; Leu or Phe was carried out by Brahmachari *et al.* (1978). Using CD, IR and NMR spectral techniques, these authors observed that the major ordered conformation is the β-turn (this conformation will be in equilibrium with the "open" disordered conformation in solution). The relative stability of the β-turn conformation was found to depend on the nature of the solvent (being more dominant in non-polar media) and more interestingly, on the nature of the X residue. In terms of their relative effectiveness in augmenting the β-turn conformation at the Pro-Gly segments, Brahmachari *et al.* (1978) found the following order: Leu > Ala > Gly, Ile > Phe. In support of the preference of the Pro-Gly segment for the β-turn conformation, Brahmachari *et al.* (1981) observed a classical type II β-turn in the crystal structure of N-Acetyl-Pro-Gly-Phe-OH using the X-ray diffraction technique (This was the first demonstration of the β-turn in a linear tripeptide in the crystalline phase).

Statistical analysis of the crystal structure data on globular proteins by Chou and Fasman (1977) had shown that Pro in the second and Gly in the third position of the β-turn have relatively very high β-turn potentials. Using this information, Brahmachari and Ananthanarayanan (1979) examined the amino acid sequence of the nascent procollagen chains and observed that the Pro-Gly

segments in these chains exist mainly in the β-turn conformation, while almost none of the Gly-Pro segments preferred this conformation. These studies also indicated that not all the Pro-Gly segments in the nascent procollagen molecule exhibit the same preference for the β-turn formation, this being governed by the nature of adjoining residues. This is in line with the observation made earlier by these authors on the tripeptides.

1.7.7. Conformational Basis for Proline Hydroxylation by Prolylhydroxylase

Based on the above conformational considerations, Brahmachari and Ananthanarayanan (1979) have proposed that the β-turn conformation, adopted by the Pro-Gly segments, in the nascent procollagen molecule (prior to proline hydroxylation) are specifically recognized by prolylhydroxylase and thus this feature is a conformational requirement for proline hydroxylation. They also have proposed that the extent of enzymatic hydroxylation of proline residues, which is known to be influenced by the nature of the residues adjoining the Pro-Gly segment (Prockop *et al.*, 1976; Bhatnagar and Rapaka, 1976) may be related to the extent of stabilization of the β-turn by these residues in a given tipeptide sequence. This hypothesis thus offered an explanation which was not available up until then from earlier studies, in conformational terms, for the specific recognition and hydroxylation of X-Pro-Gly segments in nascent procollagen.

An important corollary that stems out of the β-turn hypothesis is the conformational need for the hydroxylation of proline residues in procollagen. An examination of the crystal structure data on t-Boc-Pro-Gly-OH, N-Piv-Pro-Gly-

OH and t-Boc-Leu-Pro-Gly-OH (Benedetti, 1977) in terms of their ϕ_{Pro}, ψ_{Pro}, ϕ_{Gly} and ψ_{Gly} values was made by Brahmachari and Ananthanarayanan (1979). These authors observed that all three -Pro-Gly- segments that do not take up a β-turn conformation have their ϕ, ψ angles suitable for the extended PP-II conformation, in spite of the flexibility around the N-C^{α} and C^{α}-C' bonds of Gly residue. Similar examination of the ϕ_{Gly}, ψ_{Gly}, ϕ_{Pro} and ψ_{Pro} values for several of the available linear peptides with -Gly-Pro- sequences, namely, t-Boc-Gly-Pro-OH, t-Boc-Gly-Pro-OBzl, Cbz-Gly-Pro-OH, Cbz-Gly-Pro-Leu-OH, Cbz-Gly-Pro-Leu-Gly-OH, and O-Br-Cbz-Gly-Pro-Leu-Gly-Pro-OH (Benedetti, 1977) was also carried out by Brahmachari and Ananthanarayanan, (1979). It is observed that some of the Pro residues take up a low ψ value ($\sim$ -30°) (suitable for type I β-turn) whereas others fall in the high ψ ($\sim$ 150°) region (suitable for extended PP-II conformation). On the other hand, the Gly residues in all these peptides have relatively high ψ values ($\sim$ 170°), which are close to the value found for Gly in collagen or PP-II like extended conformation. The Pro-Gly sequences that adopt a folded β-turn conformation will have to cross a higher energy barrier of at least 4-5 K cal/mol to go over to the ϕ, ψ region of the PP-II or triple-helical native collagen. In contrast, the -Gly-Pro- and -Pro-Gly- sequences that are suitable for the extended conformation will have to cross a relatively smaller energy barrier of only 1-2 Kcal/mol to attain the ϕ, ψ required to form the triple-helical or PP-II conformation (Kolaskar *et al.*, 1975). Brahmachari and Ananthanarayanan (1979) argued that this energy barrier would be overcome by the enzymatic hydroxylation of Pro-Gly sequences in nascent procollagen to Hyp-Gly sequences, since in the latter sequences, the extended structures would be

stabilized by the H-bond involving the OH-group of the Hyp residues. [As mentioned in section 1.3.1, this intra-molecular H-bonding between the γ-OH group of Hyp (i)th residue and the carbonyl oxygen of the (i-2)th residue in the polypeptide chain has been demonstrated in poly(Hyp) by Bansal *et al.* (1979) and Torchia *et al.*, (1972).] Based on these considerations, the need for the enzymatic conversion of a Pro-Gly segment into a Hyp-Gly segment in nascent procollagen is attributable to the need for crossing the energy barrier between the flexible, disordered conformation of the nascent procollagen peptides and the rigid, extended conformation of the final product, namely, the native collagen. The conformational consequence of peptidyl proline hydroxylation may thus be viewed as the "straightening" of the original "folded" β-turn conformation into the rigid and extended structures as shown in Figure 1-6. The energy requirement for this process can be met not only from the formation of intra-chain H-bonding but also from the subsequent inter-chain H-bonding. In conformational terms, this would mean that the low ψ_{Gly} values found in folded β-turns are translated into the high ψ_{Gly} values favoured in rigid and extended conformation similar to PP-II, given the low ϕ_{Pro} value fixed at about -60°. It is obvious (Figure 1-6) that the straightening of the individual polypeptide chains is a prerequisite for the subsequent formation of the triple-helix.

Chopra and Ananthanarayanan (1982) have provided experimental data which support both of the above postulates concerning the conformational requirement and conformational consequences of proline hydroxylation using the peptides: t-Boc-Pro-Gly-Ala-OH, t-Boc-Pro-Gly-Val-OH, t-Boc-Pro-DAla-OH and t-Boc-Gly-Val-Pro-Gly-Val-OH. CD, IR and NMR spectral data showed that

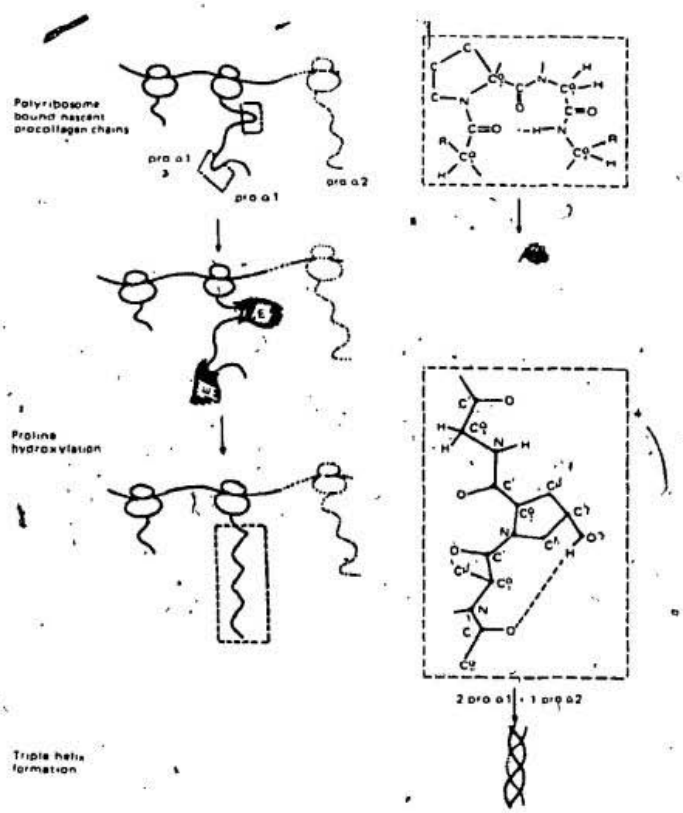


Figure 1-6: Conformational Events taking place during the Proline Hydroxylation in Collagen Biosynthesis

(The N-terminal globular extensions on the nascent chains are not shown in the figure)

From Brahmachari and Ananthanarayanan, (1979)

these peptides exist in the β-turn conformation (Ananthanarayanan and Shyamasundar, 1981; Brahmachari *et al.*, 1982). Chopra and Ananthanarayanan (1982) showed that all the peptides interacted with the enzyme by effectively competing with the standard substrate for prolylhydroxylase. Moreover, the pentapeptide t-Boc-Gly-Val-Pro-Gly-Val-OH, which acted as a substrate and also as an inhibitor to the hydroxylation of the standard substrate, was found to be mainly in the β-turn conformation, thus supporting the hypothesis that prolylhydroxylase recognizes the β-turn conformation present in Pro-Gly segments.

Chopra and Ananthanarayanan (1982) further showed that the collagen model peptide $(Pro\text{-}Pro\text{-}Gly)_{10}$ when incubated (after heat denaturation, see section 1.6.3) with the enzyme and necessary cofactors at 37 °C, was converted upon hydroxylation, from an initially non-helical conformation to the triple-helical conformation, as monitored by CD spectra and gel filtration. This interesting observation provided experimental support to the conformational hypothesis of Brahmachari and Ananthanarayanan (1979) that the straightening of the previously folded polypeptide chain occurs as a direct consequence of proline hydroxylation; this readily leads to the formation of the triple-helical conformation. In other words, triple-helix formation during collagen biosynthesis is not dependent on events distal to proline hydroxylation, unless its formation is deliberately hindered, say, by the N- or C-terminal end regions (see Prockop *et al.*, 1976; Bornstein and Traub, 1979).

Experiments on the rate of folding of the denatured $(Pro\text{-}Pro\text{-}Gly)_{10}$ and its

hydroxylated counterparts were performed by Chopra and Ananthanarayanan (1982). These studies showed that the latter polymers fold into the triple-helical conformation much faster when compared to the former. This suggests that the enzymatic proline hydroxylation in the substrates results in the enhanced rate of formation of the triple-helical conformation, indicating a hitherto unknown kinetic role for Hyp in the folding of collagen chains. According to the recent conformational energy calculations by Bansal and Ananthanarayan (1987), the Hyp-containing peptides are found to be capable of adopting a partially extended conformation which does not contain the β-bend but retains the (Hyp)OH...OC(Gly) H-bond between the $(i+2)$th Hyp and (i)th Gly. These authors propose that the partially extended conformation in the Hyp-containing peptides (which is not possible in the Pro counterparts) could serve as a kinetic intermediate on the way to forming the fully extended conformation. Because of the presence of the above intra-chain H-bond, this conformation would also serve to lock the trans-geometry at the X-Pro and X-Hyp peptide bonds, thereby enhancing the rate of their helix formation. It is known that the cis--trans isomerization around Pro-containing peptide bonds is the rate-determining step in the folding of globular proteins (Brandts *et al.*, 1975) and collagen (Bachinger *et al.*, 1978; Bruckner and Eikenberry, 1984).

In summary, the studies described above offer an explanation for the conformational criteria for, and consequences of, proline hydroxylation in collagen biosynthesis (Ananthanarayanan, 1983a). As pointed out earlier by Brahmachari and Ananthanarayanan (1979), the β-turn may serve as a general conformational requirement for several other posttranslational modifications as well. For

example, many phosphorylating enzymes (see Smith and Pease, 1980 for review) and signal sequence modification enzymes (Giam *et al.*, 1984) seem to recognize β-turn conformations in their substrate molecules for selective modification (Rose *et al.*, 1985). Recent studies by Tinker *et al.* (1986) have shown that a protein tyrosine kinase recognizes and specifically hydroxylates the tyrosine residue in a β-turn-forming tetrapeptide sequence, offering support to the above hypothesis.

1.8. Rationale and Objectives of the Thesis

In spite of the large number of studies available on proline hydroxylation in collagen as reviewed in the previous sections, there are still many aspects that need to be explored in order to understand the complex and intricate reaction carried out by prolylhydroxylase. These include:

1. Structural characteristics of, and structure-function relationship in prolylhydroxylase

2. Conformational aspects of interactions between the cofactors and cosubstrates and the enzyme with a view to understanding the mechanism of proline hydroxylation

3. Physico-chemical characterization of Pro-containing peptides in a variety of conformations (β-turn, PP-II and other conformations)

4. Study of the interactions of these peptides with prolylhydroxylase in biochemical as well as in conformational terms

These are the topics that have been chosen for study in the present thesis.

The conformational aspects of prolylhydroxylase have not yet been studied although the enzyme has been purified to homogeneity for the past several years. My approach has been to obtain pure prolylhydroxylase from chicken embryos and to characterize the enzyme in terms of the secondary and tertiary structures using CD and fluorescence spectroscopy. The same physical techniques were used to correlate the structure of the enzyme with functional aspects under a variety of conditions that can affect the structure-function relationship. The interactions between different cosubstrates and cofactors and the enzyme may lead to conformational changes which may be necessary for the function of prolylhydroxylase. These conformational changes may occur in either the secondary or the tertiary structure of the enzyme and can therefore, be conveniently monitored by CD and fluorescence techniques. The first part of the thesis is concerned with these studies.

The β-turn hypothesis has been studied so far only with a limited number of peptide models. Further verification of the hypothesis is needed for a better understanding of the substrate specificity of prolylhydroxylase, in conformational terms. In the light of the larger data base, when available, it may be possible that this hypothesis would need modification so as to account for the additional data. With this view in mind, following the purification and characterization of prolylhydroxylase (Chapters 3 and 4), an attempt has been made in the present thesis, to characterize a number of Pro-containing oligopeptides further by using CD and IR spectral techniques (Chapter 5).

The interaction of the above model peptides with prolylhydroxylase forms

another aspect of the present studies (Chapter 6). These studies are aimed at providing insights into the conformational requirements at the binding and catalytic sites of prolylhydroxylase. Detailed kinetic analysis of the peptide-enzyme interactions in terms of hydroxylation of the substrates and/or competitive inhibition of the enzyme by model peptides (with respect to the standard substrate of the enzyme) could lead to a definition of the conformational specificity of the enzyme. These studies constitute the last part of the thesis.

Other aspects related to proline hydroxylation have also been attempted for study in a limited way. These include the specificity of prolylhydroxylase with non-collagenous, biologically important proteins which will be of interest in the light of other non-collagenous, Hyp-containing proteins. It is not well-understood, as yet, whether the same protocollagen prolylhydroxylase is responsible for the formation of Hyp residues in these proteins also.

The experimental details and results of the above aspects are presented and discussed in detail in the following chapters.

Chapter 2

Materials and Methods

2.1. Materials

All the Pro-containing oligopeptides were synthesized by Dr. S. K. Attah-Poku using standard solution-phase techniques of peptide synthesis (Bodanski and Bodanski, 1984). They were checked for purity by Dr. Attah-Poku using HPLC and elemental analysis (See the above ref. for details). $(Pro\text{-}Pro\text{-}Gly)_5$ was obtained from Protein Research Foundation, Osaka, Japan. Poly(Hyp) was from Miles-Yeda chemical company. White leghorn chicken eggs were obtained from Cook's Hatchery (Truro, Nova Scotia).

The following were purchased from Sigma Chemical Company (St. Louis, Missouri, USA): Poly(Pro) samples of high and low molecular weights (M_r) (40,000 and 6,000), Trizma base, glycine, dithiothreitol, ferrous sulphate, ascorbic acid, fatty acid free bovine serum albumin, catalase, trypsin inhibitor from chicken egg white, α-ketoglutaric acid (sodium salt), Chloramine T, Folin-Ciocalteu reagent, β-mercaptoethanol and glycerol.

Ammonium sulphate, sodium chloride, sodium hydroxide, potassium dihydrogen phosphate, trichloroacetic acid and p-dimethylaminobenzaldehyde were purchased from BDH (Canada) Ltd., Dartmouth, Nova Scotia.

Chloroform, methanol, 1,4-dioxane, glacial acetic acid, hydrochloric acid and perchloric acid were obtained from Fisher Scientific Company, Dartmouth, Nova Scotia.

Cyanogen bromide-activated Sepharose and Sepharose 4B were obtained from Pharmacia, Dorval, Quebec. DE-52 ion exchange cellulose (or resin) was purchased from Mandel Scientific Company, Ville St. Pierre, Quebec. Biogel A-1.5m and Dowex 50 X-W8 were from BioRad (Canada), Mississauga, Ontario, as were electrophoresis chemicals like acrylamide, bis-acrylamide, N,N,N',N'-tetramethylethylenediamine (TEMED), ammonium persulphate, sodium dodecylsulphate and low molecular weight standards. Ultra pure urea, trifluoroethanol (Gold-labelled) and Triton X-100 were purchased from Pierce Chemical Company, Rockford, Illinois, USA.

Radioactive chemicals, namely, 2-Keto (1-^{14}C)glutarate (sodium salt), (L-Prolyl-4-^{3}H)(Pro-Pro-Gly)$_5$ were purchased from Amersham International, Oakville, Ontario. Tissue solubolizer (NCS), Omnifluor and Aquasol were from New England Nuclear, Boston, Massachusetts, USA.

2.2. Methods

2.2.1. Preparation of poly(Pro)-Agarose Affinity Column

Poly(Pro) with a M_r of 40,000 was covalently coupled to Sepharose 4B by the cyanogen bromide (CNBr) agarose activation technique. Two methods were employed:

1. Commercially available CNBr-activated Sepharose 4B was used for coupling according to manufacturer's instructions. About 15 g of freeze-dried powder was allowed to swell for 60 min in 1 mM HCl at room temperature (25 °C) and washed on a sintered (coarse) glass filter with the same solution. A total of approximately 200 ml per gram of dry gel was added in several aliquots, the supernatant being removed by suction between successive additions. One gram freeze-dried powder gives a gel volume of approximately 3.5 ml. The use of HCl preserves the activity of the reactive groups which hydrolyze at alkaline pH. The ligand solution was prepared by dissolving 500 mg of poly(Pro), in 50 ml coupling buffer-1, giving a final concentration of 10 mg/ml. The coupling buffer-1 was made up of 0.1 M sodium bicarbonate containing 0.5 M sodium chloride and the pH was adjusted to 8.3. At the end of the HCl wash, the gel was washed with coupling buffer-1 using about 5 ml of buffer per gram dry gel. The ligand solution was added immediately to the washed gel which was rendered alkaline by the coupling buffer-1. This stage (washing the gel and addition of the ligand) was completed within 40 sec to avoid any delay since reactive groups on the gel hydrolyze immediately at the alkaline pH of the coupling buffer. The gel in the poly(Pro) solution was carefully transferred to a 250 ml round-bottom flask and rotated very slowly at 4 °C for 24 h. High speeds of rotation should be avoided since the high shearing forces can damage the gel matrix.

2. CNBr-activation of Agarose and coupling of poly(Pro): This was done according to Tuderman *et al.* (1975). About 100 ml of Sepharose 4B was washed three times with one litre of water and adjusted to pH 11 and decanted to a final volume of about 200 ml. Then, 25 g of carefully ground cyanogen bromide was added and the reaction was allowed to proceed for 15-20 min in an ice bath with continuous stirring. The solution was maintained at pH 11 by frequent additions of 8 M and 1 M NaOH. The mixture was rapidly washed in a Buchner funnel with 500 ml of coupling buffer-2 made up of 0.14 M NaCl and 0.1 M sodium bicarbonate pH 8.3 and the coupling was carried out by quickly stirring 10% poly(Pro) solution (prepared as 1 g/10 ml coupling buffer-2). The washing and addition of poly(Pro) to the gel was completed within 40 s. The gel and poly(Pro) mixture was transferred to a 250 ml round-bottom flask and rotated very slowly at 4 °C for 24 h.

At the end of this 24 h period (for both the above methods), the poly(Pro) solution was removed by suction and the gel was mixed with 1 M ethanolamine pH 8.0 to block unreacted groups and kept for about 2 h at room temperature. The blocking agent was washed away with respective coupling buffer followed by 0.1 M acetate buffer at pH 4 containing 0.5 M NaCl. This was repeated a few times in alternate cycles of coupling buffer followed by acetate buffer. Alternate pH changes like these not only remove excess of blocking agent, but also the excess of ligand present due to non-specific adsorption.

Finally the gel was washed with a large volume of cold deionized water followed by the "affinity buffer" made up of 0.1 M NaCl, 0.1 M glycine, 0.1 M

Tris-HCl containing 10 μM dithiothreitol (DTT) and the pH was adjusted to 7.8 at 4 °C. The gel was equilibrated thoroughly with the same solution (about 10 column volumes). The gel was preserved with 0.02% sodium azide, a bacteriostatic agent, until ready to use.

The efficiency of coupling was checked by the following methods. A small aliquot of the gel was washed with a large volume of water to remove the glycine containing buffer and hydrolyzed in 6N HCl at 120 °C for 24 h. The hydrolysate was analyzed for proline using an amino acid analyzer at the local amino acid analysis facility.

In another method, the amount of poly(Pro) bound to agarose was measured by monitoring the difference in the absorbance of the ligand solution, at 230 nm, before and after the process of coupling reaction.

2.2.2. Purification and Characterization of Prolylhydroxylase

Prolylhydroxylase was isolated from fertilized chicken embryos according to the procedure of Tuderman *et al.* (1975) as modified by Kedersha and Berg, (1981). The details are given below.

2.2.2.1. Preparation of Embryo Homogenate

Medium-to-large sized fertilized white leghorn chicken eggs were incubated for 13-14 days in an incubator (model 1202 from G.Q.F Manufacturers, Savannah, Georgia, USA) equipped with autorotatory and thermostatic controls. The humidity of the chamber and air circulation were maintained throughout the incubation period. A constant temperature of 100 + 1° F was maintained. The

eggs were rotated 3-4 times a day automatically and were candled on alternate days for signs of viability and growth.

The eggs were opened on the 13th or 14th day and the embryos were quickly separated from the rest of the membranes and transferred to an ice-chilled beaker. The embryos were homogenized in batches of 250 g with an equal volume of homogenizing medium containing 0.1 M glycine, 0.1 M KCl, 0.1% (w/v) Triton-X, 0.03% (w/v) trypsin inhibitor (from chicken egg white) and adjusted at 4 °C, to pH 7.8 with HCl. The homogenization was carried out in a Waring blender in 3 cycles of low, medium and full speeds for 1 min at each speed and allowing 1 min for cooling between the cycles. The homogenate was allowed to stand for 30 min with occasional stirring and then centrifuged at [illegible]00 x g for 45 min and the supernatant was collected.

2.2.2.2. Ammonium Sulphate Fractionation

Solid ammonium sulphate was slowly stirred into the supernatant to a final concentration of 30% (164 g/L) at 4 °C and allowed to stand for 1 h. After centrifugation at 15,000 x g for 1 h, the pellet was discarded and solid ammonium sulphate was slowly stirred again into the supernatant to a final concentration of 60% saturation (184 g/L). After centrifugation at 15000 x g for 1 h, the pellet was dissolved in the affinity buffer containing 0.1 M Tris-HCl, 0.1 M NaCl, 0.1 M glycine and 0.01 mM DT[illegible]d adjusted at 4 °C to pH 7.8. The dissolved pellet was dialyzed for 4 h against 16 L of affinity buffer and for another 12-16 h after changing the buffer.

2.2.2.3. Affinity Chromatography

The dialyzed ammonium sulphate precipitate was centrifuged at 15000 x g for 30 min and the protein concentration of the supernatant was measured according to Lowry *et al.* (1951) or the BioRad Protein Assay (1977). The concentration of the protein was adjusted to 10 mg/ml with affinity buffer. The sample (about 150 ml) was then mixed with the washed poly(Pro)-Sepharose (100 ml) prepared as described above and binding of the enzyme was allowed to take place with occasional stirring for about 8 h at 4 °C. Initial and subsequent samples were removed to monitor the extent of binding at different time intervals by assaying the activity remaining in the supernatant. At the end of the binding period, the gel was washed with large volumes of affinity buffer on a sintered glass funnel in the cold, until the absorbance of the eluent was less than 0.3 at 230 nm using 1 cm cell. The gel was packed into a column at 60 ml/h flow rate. The enzyme was eluted with 20 ml (3 mg/ml) of poly(Pro) solution in the same buffer at a flow rate of 50 ml/h. The elution was then continued by one column volume of affinity buffer and 4 ml fractions were collected. The poly(Pro) used for elution has a M_r of about 6,000. The size of the poly(Pro) is important, in that a peptide of M_r more than 8,000 cannot be completely separated from the enzyme in the later stages of purification.

The absorbance of the fractions eluted with poly(Pro) was monitored at both 280 and 230 nm. The enzyme-poly(Pro) complex eluted as a symmetrical peak having an absorbance at both 280 and 230 nm wavelengths. The fractionscomprising the peak were combined together and concentrated to about 5 ml in an Amicon ultrafiltration cell with YM-30 membrane (molecular weight cut-

off: 30,000) under 20 psi of nitrogen gas (N_2). The ultrafiltration cell was washed twice with 0.5 ml of affinity buffer and these washes were pooled with the enzyme-poly(Pro) complex concentrated earlier. The ultrafiltration step was also useful in removing the bulk of the poly(Pro).

2.2.2.4. DEAE Ion Exchange Chromatography

The enzyme-poly(Pro) complex (about 8 ml) was dialyzed overnight at 4 °C, against 2 L of the "ion-exchange buffer" containing 0.05 M Tris-HCl, 0.1 M glycine, 0.05 M NaCl and 0.01 M DTT and adjusted to pH 7.4 at 4 °C. This buffer was also used to equilibrate the DEAE ion-exchange resin. The dialyzed material was applied at a flow rate of 40 ml/h to the pre-equilibrated DEAE ion-exchange column (1 x 18 cm) and washed with about 50 ml of the same buffer. The enzyme, being negatively charged at this pH, binds to the resin whereas the uncharged poly(Pro) is washed off within one column volume of the buffer. The bound enzyme was eluted by applying a 300 ml linear gradient of increasing NaCl concentration (0.05 M - 0.35 M) in the same buffer. The high ionic strength buffer is the limit buffer and is referred to as the "gradient buffer". Fractions of 4.5 ml were collected and the absorbance at both 280 and 230 nm was monitored. Two peaks were eluted usually with significant resolution during this gradient elution. Prolylhydroxylase activity was assayed (see 2.2.3.2 and 2.2.3.3) and the fractions contained in the second peak showed the activity. These fractions were pooled and concentrated to 2 ml in an Amicon ultrafiltration cell through a YM-10 (molecular weight cut-off 10,000) membrane at 25 psi of N_2 pressure. At the end of the concentration, the cell was washed with two 0.5 ml aliquots of buffer and these washes were pooled with the concentrate. The concentration was

monitored by the increase in the absorbance of the sample at 280 and 230 nm during the course of concentration.

2.2(2.5. Gel Filtration Chromatography

The concentrated enzyme sample was centrifuged at 2000 x g for 10 min and the clear supernatant was mixed with a few drops of glycerol and applied to a pre-equilibrated Biogel A-1.5m column (1.5 x 85 cm) and eluted with the affinity buffer at 10 ml/h. Fractions of 3 ml were collected and their absorbance was measured at 280 and 230 nm. The enzymatic activity was assayed and the fractions containing both absorbance and the activity were pooled and concentrated by ultrafiltration (through YM-30) as described above and then centrifuged at 2000 x g for 5 min. The clear supernatant was stored in small 0.2-0.5 ml aliquots frozen in liquid N_2 until ready to use. Storage in this manner was found to prevent loss of activity due to repeated thawing of the enzyme preparation.

2.2.2.6. Protein Estimation

The amount of protein present in samples at various stages of enzyme purification was estimated by either Lowry's method (1951) or by BioRad Protein Assay (1977). For pure prolylhydroxylase, the concentration was estimated using an extinction coefficient of 1.54 at 280 nm and 7.73 at 230 nm for a 1 mg/ml solution (Kivirikko and Myllyla, 1982).

2.2.3. Characterization of Prolylhydroxylase

2.2.3.1. Assays of Enzymatic Activity

The activity of the enzyme was assayed by one or more of the following methods : $^{14}CO_2$ release method (Rhoads and Udenfriend, 1970; Kivirikko *et al.*, 1972), colorimetric estimation of hydroxyproline (Hyp) formation (Edwards and O'Brien, 1980) or the tritiated water release method (Peterkofsky and Diblasio, 1973). Both $^{14}CO_2$ release method and colorimetric method were used routinely for the estimation of Hyp produced in the enzyme assays. Tritiated water release method was used only in the inhibition experiments. All the three methods were standardized for the optimal assay conditions, as described below. The sensitivity of the $^{14}CO_2$ method is about 2 nmol (200-300 cpm above the background value of 40-50 cpm) while that of the colorimetric method is about 8 nmol of Hyp produced. In the case of tritiated water assay method, the sensitivity of the method is about 10 nmol of the product formed (100-150 cpm above the background value). In the case of $^{14}CO_2$ release method, the sample cpms were corrected for the cpms due to the uncoupled decarboxylation under the same conditions. Similarly, any non-specific tritium release was checked and corrected from the sample values in the tritiated water release assay. The specificity of the colorimetric method was similarly checked by treating the peptide substrate in the absence and presence of the enzyme under identical reaction conditions.

In the earlier stages of purification, the enzymatic activity is defined in arbitrary units of μ moles of Hyp formed by one ml of enzyme preparation, in one hour at 37 °C under the conditions specified below. The specific activity of the

pure enzyme is defined in terms of the μ moles of Hyp produced in one hour by one milligram of enzyme protein at 37 °C under the specified conditions.

2.2.3.2. $^{14}CO_2$ Release Method

This method is based on the stoicheiometric decarboxylation of the α-KG during the hydroxylation of the peptide substrate by prolylhydroxylase. In most experiments, the activity was assayed in 25 ml Erlenmeyer flask with a final reaction volume of 1 ml. The reaction components were added in the following specific order, to the following final concentrations: 0.05 M Tris-HCl pH 7.4, 2 mg/ml bovine serum albumin (BSA), 0.1 mg/ml catalase, 0.1 M DTT, 2 mM ascorbate, 5-10 μg/ml prolylhydroxylase, 0.1 mM $FeSO_4$, 2 mM ^{14}C-1 α-KG with a specific activity of 0.1 μC/ml (2.1 X 10^5 dpm) and finally 0.5 mg/ml of $(Pro\text{-}Pro\text{-}Gly)_5.4\ H_2O$ as the substrate . The reaction mixture was pre-equilibrated at 37.5 °C for 5-8 min before the addition of the peptide substrate to start the the reaction. The flasks were closed with rubber stoppers fitted with plastic centre wells (from Mandel scientific company); these wells contained 0.3 ml of NCS (Amersham) solution which absorbs released radioactive carbon dioxide. The reaction mixtures were incubated for 15 min and arrested by injecting1 ml of 1 M potassium phosphate buffer, pH 4.0 into the flask, care being taken not to spill into the centre well. The released $^{14}CO_2$ was trapped into the NCS solution for about 4 hours and then the centre wells were cut out and transferred to the scintillation vials and counted in Omnifluor-toluene scintillation fluid. A Beckman liquid scintillation counter with an 80% efficiency of counting for ^{14}C was used.

The peptide substrate was thermally denatured by heating to 100 °C for 5 min and quenching to 0 °C immediately before adding to the reaction mixture. Both DTT and $FeSO_4$ were dissolved in water, bubbled with dry N_2 for 20-30 min to prevent oxidation of these chemicals.

2.2.3.3. Colorimetric Estimation of Hydroxyproline

The colorimetric estimation of Hyp is according to the method of Edwards and O'Brien (1980). This method is based on the oxidation of Hyp by Chloramine T (sodium N-chloro p-toluene sulphonamide) in citrate-acetate buffer near neutrality to a pyrrolic compound. This compound reacts with Ehrlich's reagent (p-dimethylamino benzaldehyde) in strong perchloric acid to give a red chromophore which has an absorption maximum at 550 nm. The reaction was carried out exactly as above for the $^{14}CO_2$ release method except for the use of unlabelled α-KG. The reaction was arrested with 1.0 ml of conc. HCl (12 N) and the samples were transferred to hydrolysis tubes and hydrolyzed at 120 °C for 24 h under vacuum. The hydrolyzed samples were evaporated to dryness under vacuum at 70 °C to remove HCl and the residue was dissolved in 0.5 M citrate-acetate buffer pH 6.5. Two ml of the sample were reacted with 1.5 ml of 0.05 M Chloramine T for 25 min at 25 °C. Ehrlich's reagent (1.5 ml) was added and samples were incubated at 60 °C in a shaking water bath for 15 min for the development of the red chromophore. At the end of this period, the samples were cooled to room temperature and the absorbance of the samples was read at 550 nm in a Pye-Unicam PU 8800 UV-Vis spectrophotometer with 1 cm pathlength glass cuvettes. A calibration curve was constructed for each experiment using known amounts (1-8 μg) of Hyp. The amount of Hyp produced in the samples was estimated from the standard curve.

2.2.3.4. Tritiated Water Release Assay

When tritiated peptide substrate was used in the reaction, the assay was performed as described by Peterkofsky and Diblasio (1973). The reaction was carried out as described earlier except that the standard substrate was mixed with radioactive (L-Prolyl-4-^{3}H) $(Pro\text{-}Pro\text{-}Gly)_5$ (0.1 μC/mg $(Pro\text{-}Pro\text{-}Gly)_5$). The total reaction volume was 0.5 ml. The reaction was arrested with 0.1 ml of 10% TCA (trichloroacetic acid). The samples were allowed to stand for 20 min for precipitation of the proteins and then centrifuged at 2500 x g for 10 min to separate the precipitate. The supernatant was transferred into another small vial; the precipitate was washed with another 0.5 ml aliquot of 10% TCA, vortexed and centrifuged again. The supernatant was added to the earlier one. The tritiated water was separated from the peptide substrate by chromatography on Dowex columns. Small (1 X 4 cm) disposable Dowex 50X-W8 (H^+ form, precycled) columns were prepared in 5 ml syringes and washed thoroughly with deionized water. The above supernatants were passed through separate columns and the eluants were collected directly into the scintillation vials. The column was washed again with two 1 ml aliquots of deionized water. The eluants were counted for radioactivity in Aquasol II (Amersham) in a Beckman liquid scintillation counter with a counting efficiency of 45-50% for tritium.

2.2.3.5. Effect of Temperature on the Activity of the Enzyme

The enzymatic reaction was carried out at different fixed temperatures between 10-90 °C and the activity of prolylhydroxylase at these temperatures was assayed.

2.2.3.6. Molecular Weight Determination

The final purification step i.e the gel filtration chromatography on a calibrated Biogel A-1.5m column gives an idea about the purity as well as molecular weight of prolylhydroxylase. The enzyme was eluted in a single symmetrical peak, at an elution volume corresponding to the molecular weight of the native enzyme.

2.2.3.7. Subunit Structure of Prolylhydroxylase

The subunit structure and the molecular weight of these subunits were confirmed by polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulphate (SDS-PAGE). The discontinuous system of Laemmli (1970) was used with a 12% separating gel prepared in 0.375 M Tris-HCl and 0.1% SDS, pH 8.8 and a 3.5% stacking gel prepared in 0.125 M Tris-HCl and 0.1% SDS, pH 6.8. Gel slabs of 14 cm X 8 cm X 0.5 cm dimensions were used. The electrophoresis apparatus is from Pharmacia, Model GE 2/4 LS. The electrophoresis buffer had the composition of 0.05 M Tris-HCl, 0.384 M glycine and 0.1% SDS, pH 8.3. The sample was prepared by heating 20-40 μg of enzyme at 100 °C for 5 min in about 25 μl of the sample buffer made up of 0.01 M Tris-HCl, 0.001 M EDTA, 1% SDS and 5% β-mercaptoethanol and 0.01% bromophenol blue. After the samples were applied to the gel, about 70 V was used until the sample gets concentrated to a thin disc in the stacking gel and 140 V was used throughout the separation. 6 to

8 hours of electrophoresis was usually sufficient for good resolution. The staining was done overnight in Coomassie Blue G-250 in 35% perchloric acid at room temperature. Destaining was carried out in 7% acetic acid with frequent changes of the solution until the background became very clear.

2.2.4. Effect of Variables on the Initial Velocity of Prolylhydroxylase Reaction

2.2.4.1. Effect of Time

The reaction mixture containing all the components was incubated for different (0-60 min) periods of time to check the linearity of the reaction under the conditions specified. In the case of the $^{14}CO_2$ method, the control flasks without the added substrate were also incubated to the same time periods as were the tests to correct for the uncoupled decarboxylation taking place during these time periods. At the end of each incubation period, the reaction was arrested and assayed as usual.

2.2.4.2. Effect of Enzyme Concentration

The reaction mixture containing all the components was incubated with different amounts of the enzyme (0-40 μg/ml) for a specified time to check the linearity of the initial velocity of the reaction, with respect to increasing amounts of the enzyme.

2.2.4.3. Effect of Substrate Concentration

The enzymatic reaction was carried out with varying amounts of the substrate $(Pro\text{-}Pro\text{-}Gly)_5$ (0-1.5 mg/ml) in the presence of all other reaction components specified. About 5-10 μg of enzyme was used. The time of incubation at each concentration of the substrate was adjusted to ensure that the reaction under these conditions remained linear.

2.2.4.4. Effect of α-KG Concentration

The enzymatic reaction was carried out with varying concentrations (0-5 mM) of α-KG with all other reaction components fixed as specified above and assayed as usual.

2.2.4.5. Effect of $FeSO_4$ Concentration

The enzymatic reaction was carried out with varying concentrations of $FeSO_4$ (0-0.2 mM) and all other reaction components fixed as specified above.

2.2.5. Circular Dichroism (CD) Measurements of Prolylhydroxylase

The secondary and tertiary structures of prolylhydroxylase were examined using circular dichroism (CD) spectroscopy. CD spectra were recorded using a Jasco J-500N spectropolarimeter equipped with the DP-500 N data processor. Water-jacketed quartz cells of 0.1-5 mm pathlength were used. The temperature was controlled with a Lauda RC 6 Model circulating water bath equipped with refrigeration and was measured with a thermistor probe (inserted at the outlet) to an accuracy of ± 0.2 °C. An enzyme concentration of about 0.5 mg/ml was used for far-ultra-violet (far-UV) CD spectra (250-185 nm region) and about 1-1.5 mg/ml was used for near-UV spectra (320-250 nm). The enzyme solution was

filtered through a Millipore filter (0.45 μm) to remove any particulate aggregates or impurities. When CD measurements were carried out at 25 °C or higher temperatures, the enzyme solution was degassed under vacuum, to prevent the formation of air bubbles at elevated temperatures. The ellipticity $[\theta]$, in units of deg.cm^{2}dmol^{-1} was calculated using a mean residue weight (MRW) of 111 for prolylhydroxylase as obtained from the amino acid composition (Berg *et al.*, 1979). The following formula was used for the calculation of $[\theta]_{MRW}$:

$$[\theta]_{MRW} = \frac{\theta \times MRW}{10 . c . l}$$

Where θ is the observed ellipticity in degrees, c is the concentration of the enzyme in g/ml and l is the pathlength in cm. The CD spectrum of the enzyme was scanned at least 5 or 6 times, each time the spectrum was repeated at least 8 times and electronically averaged using the data processor. The spectrum was corrected for the contribution from the solvent by electronically subtracting the CD spectrum of the solvent from the sample spectrum. The effect of temperature on the native conformation of the enzyme was also monitored by CD. The temperature was varied between 0 and 90 °C in steps of 2.5-5.0 °C and with an accuracy of $\pm$0.2 °C. The possible conformational changes induced in the secondary structure of prolylhydroxylase upon the binding of substrates, cosubstrates and cofactors were studied using the far-UV CD spectra in the region of 250-185 nm. An enzyme concentration of 0.7 mg/ml was used and these interactions were monitored at 37 °C in water-jacketed cells of 0.1 mm pathlength. The following concentrations of the substrate and cosubstrates were

utilized : $(Pro\text{-}Pro\text{-}Gly)_5$ between 0.1-1.0 mg/ml; α-KG between 0.2-2.0 mM; $FeSO_4$ between 0.05-1.2 mM and ascorbate in the range of 0.5-2.5 mM. CD spectra were corrected for the contributions from the ligands at each concentration. This was done by subtracting the spectrum of the ligand from the spectrum of the enzyme + ligand together.

2.2.6. Fluorescence Measurements of Prolylhydroxylase

The emission characteristics of prolylhydroxylase have been determined using fluorescence spectroscopy. The spectra were recorded in a Shimadzu spectrofluorophotometer Model RF-540 equipped with a data processor recorder Shimadzu DR-3. A microcell with about 0.3 ml minimum volume and 1 cm pathlength was used. The temperature was controlled using the special cell holder and circulating water bath to an accuracy of ± 0.5 °C. The effect of enzyme concentration on its emission properties was examined and an enzyme concentration of about 0.01-0.02 mg/ml was found to give optimal fluorescence, with intensity linearly proportional to the enzyme concentration. This concentration (0.02 mg/ml) was subsequently used in all other experiments. The enzyme solution was filtered through a Millipore filter (0.45 μm) to remove any particulate aggregates or impurities. A slit width of 5 nm was used for both excitation and emission beams. The excitation spectrum was scanned and the excitation maximum was found around 286 nm which was used in subsequent experiments. The emission range was between 310-450 nm and the spectra were corrected for solvent emission but not corrected for quantum yield. The fluorescence emission spectra of prolylhydroxylase were recorded at different fixed temperatures between 0 and 90 °C to check the effect of temperature on the

emission properties of prolylhydroxylase. The enzyme solution was degassed under vacuum in order to prevent the formation of air bubbles at elevated temperatures.

The effects of the binding of cosubstrates and cofactors to prolylhydroxylase were studied using fluorescence spectroscopy. During these studies, the excitation wavelength was set at 286 nm and the emission spectra were recorded in the presence of 0-5 mM α-KG or 0-1 mM $FeSO_4$, at different fixed concentrations of these cofactors. An enzyme concentration of 0.02 mg/ml was used and the final total volume was 0.3 ml. The interactions were studied at 37 °C. The spectra were corrected for the contributions of α-KG or $FeSO_4$ to the sample spectrum.

2.2.7. Studies on Synthetic Peptides

2.2.7.1. Infrared (IR) Spectroscopy

The IR spectra of the Pro-containing peptides were recorded in a Perkin-Elmer ratio-recording double beam IR spectrophotometer, Model 983G equipped with microprocessor control. The solvent used was chloroform ($CHCl_3$) which was purified to remove alcohol present as a stabilizer (Vogel, 1957) and distilled twice prior to use. The peptides were dried for 48-72 hours under vacuum before making up the solutions. The peptide concentration was 1-5 mg/ml and BaF_2 cells with a 1 cm pathlength were used. The spectra were corrected for the $CHCl_3$ baseline. The solutions were filtered through 0.45 μm Millipore teflon filters to remove any particulate impurities.

2.2.7.2. CD Spectroscopy

The conformation of the peptides was studied using CD spectra. Water-jacketed quartz cells of 0.01 and 0.1 cm pathlengths were used. The peptide conformation was studied in three solvents namely, trifluoroethanol (TFE), water and 1,4-dioxane. Spectra were recorded at 1 ± 0.5 °C in the case of TFE and water solutions and at 25 ± 0.5 °C in the case of 1,4-dioxane (dioxane freezes at 11.7 °C). The temperature was controlled with a circulating water bath containing 50% ethylene glycol. The peptide concentration was 1-2 mg/ml. Solutions were filtered through a Millipore teflon microfilter (0.45μm) to remove any particulate impurities. Doubly distilled and deionized water and redistilled TFE and dioxane were used. The CD spectrum of a given peptide in a given solvent was recorded in duplicate or triplicate, each time the spectrum was repeated at least 8 times and electronically averaged using the data processor. The spectrum was corrected for the contribution from the solvent by electronically subtracting the CD spectrum of the solvent from the sample spectrum. The ellipticity $[\theta]$ in deg.cm^{2}dmol^{-1} was calculated per peptide bond, using the appropriate MRW for each peptide.

2.2.8. Interaction of Prolylhydroxylase with Pro-containing Peptides

2.2.8.1. Hydroxylation of Synthetic Peptides

The following peptides were studied for their ability to act as substrates for prolylhydroxylase by undergoing hydroxylation :

1. t-Boc-Pro-Pro-Gly-NHCH$_3$ (Boc-PPG-NHCH$_3$)
2. t-Boc-Pro-Pro-Gly-Pro-OH (Boc-PPGP-OH)
3. t-Boc-Pro-Pro-Gly-Pro-NHCH$_3$ (Boc-PPGP-NHCH$_3$)
4. t-Boc-Pro-Pro-Gly-Pro-Pro-OH (Boc-PPGPP-OH)

5. t-Boc-Pro-Pro-Ala-Pro-OH (Boc-PPAP-OH)
6. t-Boc-Pro-Pro-Gln-Pro-OMe (Boc-PPQP-OMe)
7. t-Boc-Val-Pro-Gly-Val-OH (Boc-VPGV-OH)
8. t-Boc-Gly-Val-Pro-Gly-Val-OH (Boc-GVPGV-OH)

The abbreviation of the peptides, in one-letter code for amino acids is given in parentheses. t-Boc- is the abbreviation for tertiary butyloxy carbonyl group and OMe is for O-methyl ester. The hydroxylation studies were carried out in two stages. In the first stage, all the above peptides were treated with the enzyme under standard reaction conditions to check whether they underwent the hydroxylation by both $^{14}CO_2$ release method and chemical estimation of Hyp method (see 2.2.3.2 and 2.2.3.3.). The concentration of these peptides were 20 mM when compared to the 0.35 mM concentration of the standard substrate, $(Pro\text{-}Pro\text{-}Gly)_5$ which was considered arbitrarily to be 100% hydroxylated in order to serve as the reference. After the initial screening of the peptides for their ability to act as substrates for prolylhydroxylase, detailed kinetic studies were carried out. During these studies, the peptides were used in different fixed concentrations in the range of 5-40 mM.

2.2.8.2. Inhibition by Synthetic Peptides

In kinetic experiments designed to test the competitive nature of the inhibitory effect of the Pro-containing peptides on the hydroxylation of the standard-substrate by prolylhydroxylase, the following peptides were chosen as the representative examples :

1. t-Boc-Pro-Pro-Gly-Pro-$NHCH_3$

2. t-Boc-Gly-Val-Pro-Gly-Val-OH

To serve as a standard for inhibition, poly(Pro) with a M_r of about 6,000 was used. The radioactive standard peptide $(Pro\text{-}Pro\text{-}Gly)_5$, tritiated at the fourth carbon ($C\gamma$) atom of the proline ring was used and the tritium release assay procedure (2.2.2.4) was employed. The inhibitory peptides were used at a concentration of 10 mM and 20 mM in the case of Boc-PPGP-$NHCH_3$ and at 5 mM and 10 mM in the case of Boc-GVPGV-OH. The labelled peptide was diluted with the unlabelled peptide and the substrate concentration was varied between 0 and 1 mM. The time of the reaction and enzyme concentration were adjusted to give an optimal product formation at a linear rate at all concentrations of the substrate. The results were analyzed using the Lineweaver-Burk or double reciprocal plots (Lineweaver and Burk, 1934) as well as Dixon plots (Dixon, 1953) to determine the type of inhibition.

Chapter 3

Isolation and Characterization of Prolylhydroxylase

3.1. Purification of Prolylhydroxylase

As described in chapters 1 and 2, prolylhydroxylase was purified from 13-14 day old chicken embryos using an affinity procedure after ammonium sulphate fractionation. It was further purified on DEAE-cellulose ion-exchange chromatography and finally by gel filtration chromatography on a Biogel A-1.5m column. The purification procedure is schematically represented as a flow diagram in Figure 3-1. The protein precipitating between 30-60% ammonium sulphate contained approximately 60-70% of the total prolylhydroxylase activity present in the chicken embryo homogenate (see Table 3-1 on page 86).

3.1.1. Affinity Chromatography

The detailed procedure was presented earlier in Chapter 2 (2.2.1.3). Briefly, this involves affinity binding of the enzyme in the ammonium sulphate precipitate onto the "affinity gel" in which poly(Pro) (M_r of 40,000), a potent competitive inhibitor of the enzyme was covalently linked to an agarose matrix. About 100 ml of gel was used and the binding was allowed to take place as a batch procedure in a beaker rather than applying the ammonium sulphate precipitate fraction through the column containing the affinity gel. Assays of enzymatic

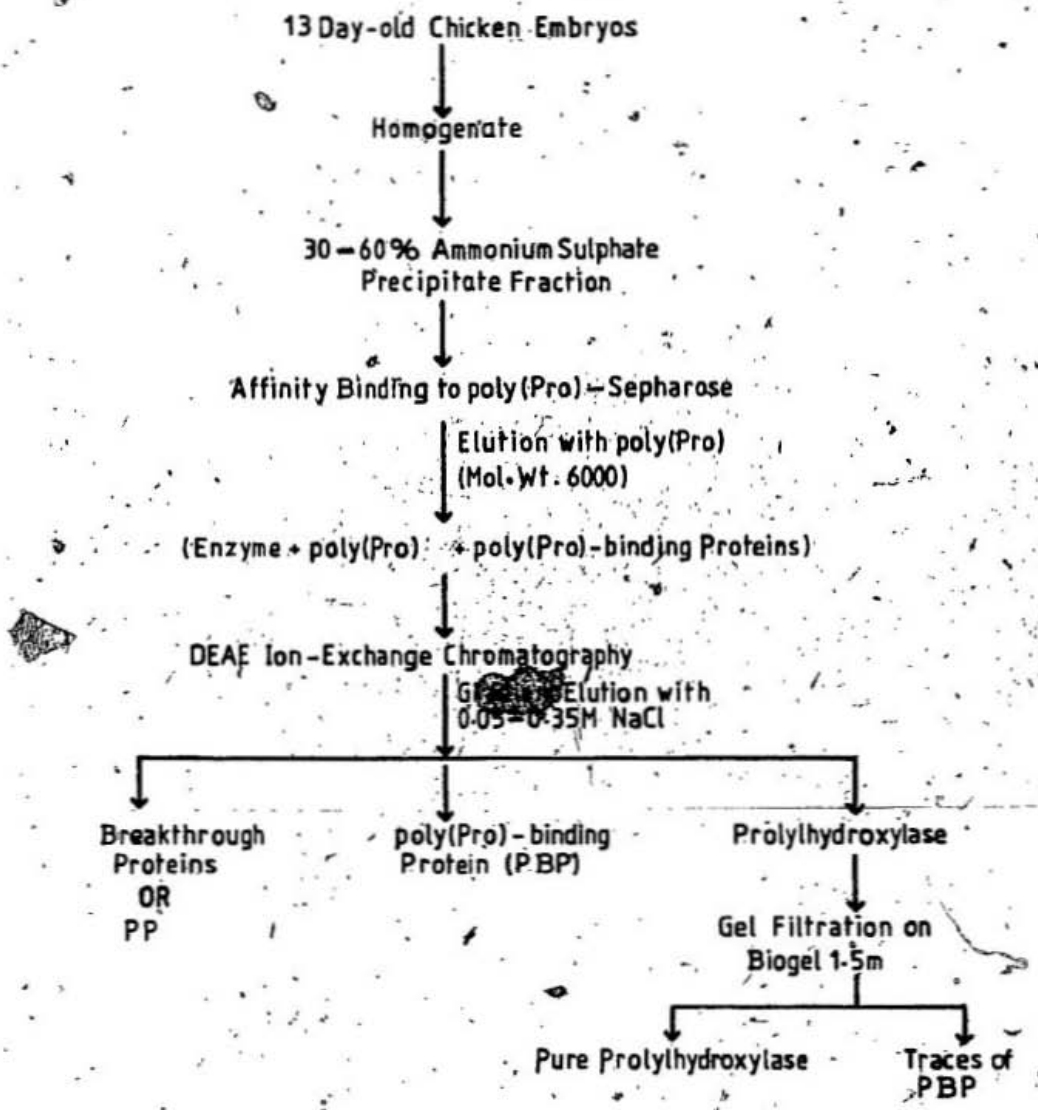


Figure 3-1: Procedure for the Purification of Prolylhydroxylase

activity in the fractions of the supernatant from the affinity binding mixture, at different time intervals, indicated that about 90-95% of the enzyme was bound to the gel in about 6-8 hours at 4 °C. The batch method has been proved to be time saving. On the other hand, when the ammonium sulphate fraction was applied to a column containing the affinity gel it took 20-25 hours to pass all the solution at a flow rate of 50-60 ml/h. Earlier studies by Kedersha and Berg (1981) also reported about 20 h for the application of the ammonium sulphate fraction onto the gel in a column. The binding of the enzyme in a batch procedure also circumvents the problems like clogged columns and slower flow rates due to the viscous nature of the ammonium sulphate fraction at low temperatures.

The elution profile of the enzyme during affinity chromatography is presented in Figure 3-2. As seen from the figure, the enzyme-inhibitor complex eluted in a single symmetrical peak between the elution volume of 40-60 ml. indicating the completion of elution. The absorbance monitored at 280 nm, at the peak height was about 2 absorbance units, when 300 chicken embryos were used as the starting material. Since the enzyme is in the presence of a potent competitive inhibitor, the enzymatic activity at this stage could not be used to monitor the amount of enzyme obtained. The entire peak containing the enzyme-poly(Pro) complex was pooled and concentrated.

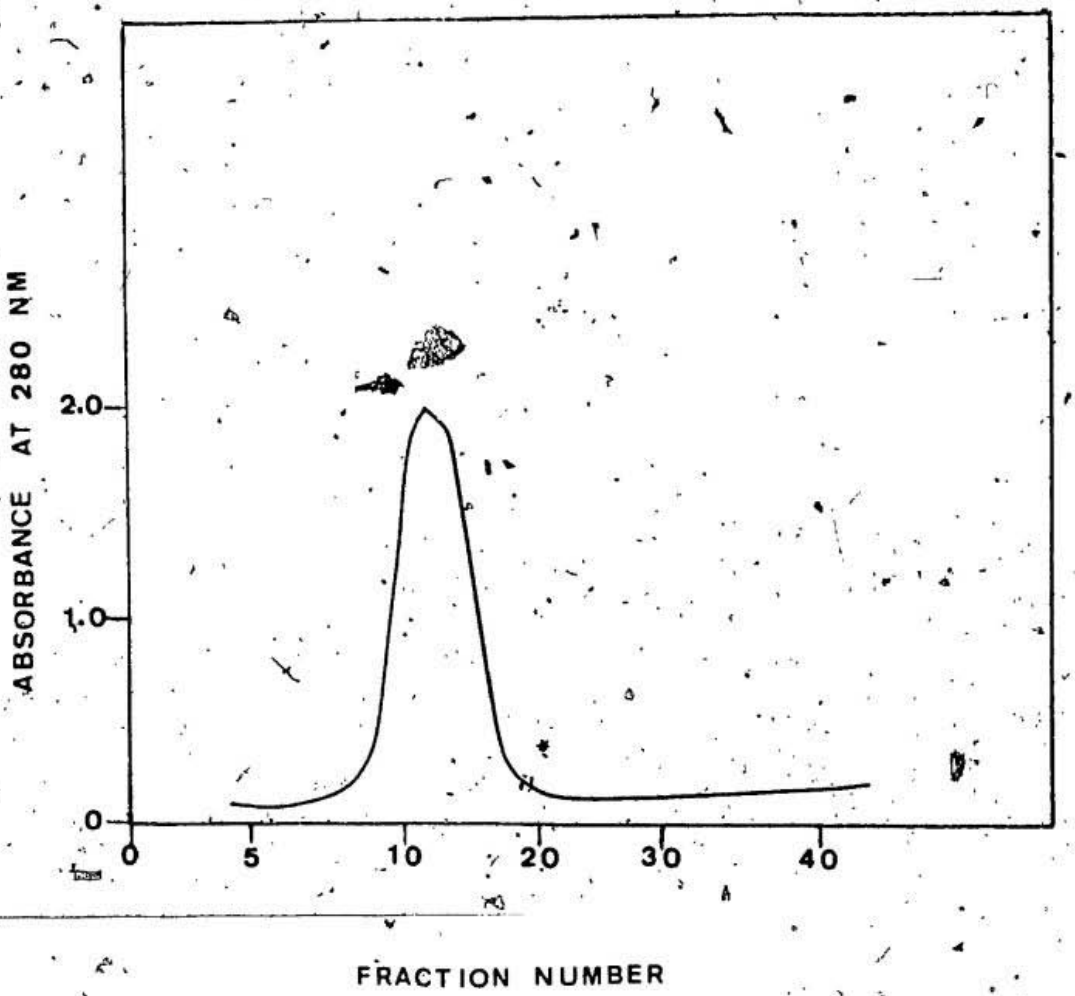


Figure 3-2: Elution of the Bound Protein from poly(Pro) Affinity Column Prolylhydroxylase was eluted with a solution of 3 mg/ml poly(Pro) (molecular weight of 6,000) dissolved in affinity buffer. 20 ml of poly(Pro) was used to elute the enzyme followed by the affinity buffer, both at a flow rate of about 50 ml/h. Fractions of 4 ml were collected.

3.1.2. Ion-exchange Chromatography

The elution profile of the enzyme during ion-exchange chromatography is presented in Figure 3-3 which reveals a few breakthrough fractions containing mainly the neutral, unbound poly(Pro). Since these fractions exhibit absorbance mainly at 230 nm (>3.0 units) and very little absorbance at 280 nm, it is unlikely to contain any proteins that bound to the poly(Pro) affinity gel. These fractions were washed off the ion-exchange column during the initial washing, before the salt gradient was applied.

Application of the sodium chloride gradient results in the elution of two peaks usually with good resolution. Peak I corresponds to a single protein, earlier referred to as poly(Pro)-binding-protein (abbreviated as PBP) by Kedersha and Berg (1981). Recently, this protein has been shown to be profilactin, a complex between profilin and actin (Tanaka and Shibata, 1985). Assay of enzymatic activity on different fractions indicated that the peak II contained 100% of the prolylhydroxylase activity eluted from the column (Figure 3-3). The fractions with prolylhydroxylase activity were pooled together and concentrated to be further purified on a Biogel A-1.5m column.

3.1.3. Gel Filtration Chromatography

The enzyme is usually separated well from PBP during DEAE-cellulose ion-exchange step. However, any minor contamination of the enzyme by PBP is further removed by gel filtration on a Biogel A-1.5m column. The elution profile is presented in Figure 3-4. The enzyme was eluted as a single symmetrical peak between 60-90 ml elution volume and the fractions under the peak exhibited both

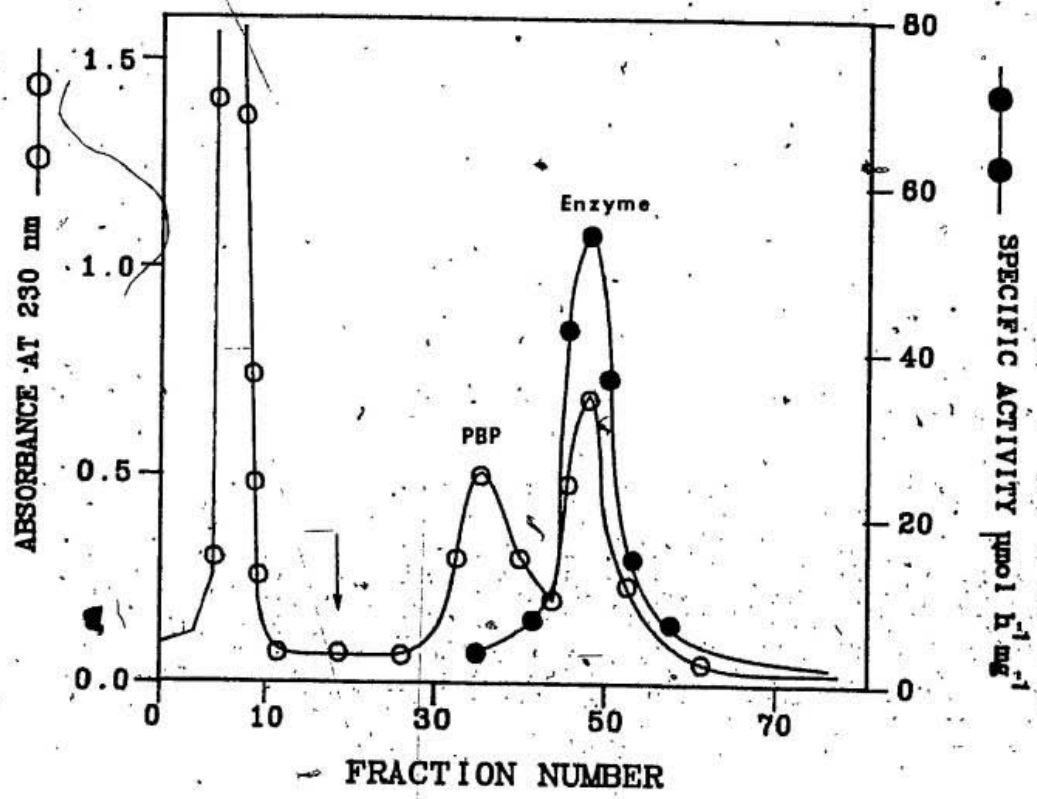


Figure 3-3: Purification of Prolylhydroxylase by DEAE Chromatography

About 15 ml of dialyzed enzyme-poly(Pro) complex was applied to the ion-exchange (1x18 cm) column at a flow rate of about 40 ml/h. The column was washed with 50 ml of ion-exchange buffer and then eluted with 300 ml linear gradient of 0.05-0.35 M NaCl in ion-exchange buffer, pH 7.4. Fractions of 4 ml were collected. The arrow indicates the start of the gradient. Poly(Pro)-binding protein (PBP) and the enzyme peaks are indicated.

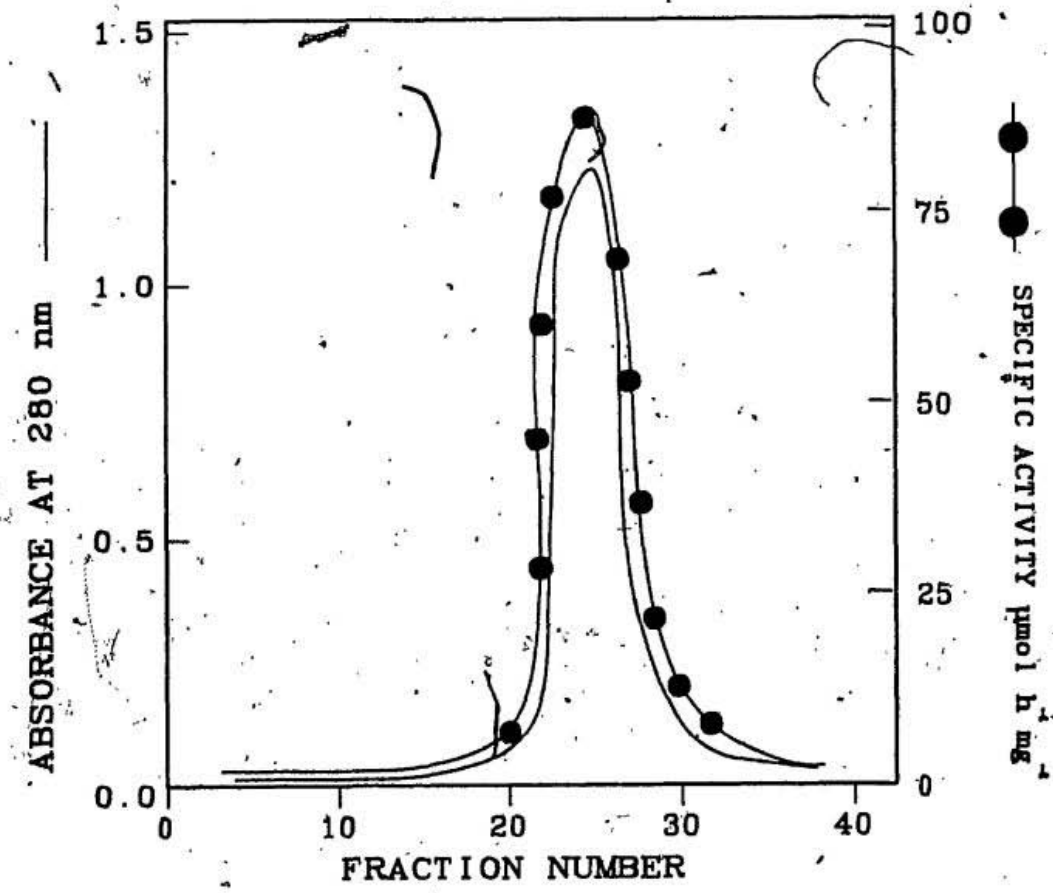


Figure 3-4: Purification of Prolylhydroxylase on Biogel A-1.5m Column Pooled and concentrated enzyme peak from DEAE chromatography (3 ml) was applied to Biogel A-1.5m gel filtration column and eluted with affinity buffer at a flow rate of 10 ml/h. Fractions of 3 ml were collected. The solid line indicates absorbance at 280 nm while the line with closed circles indicates specific activity of the enzyme in μmol/h/mg.

absorbance and enzymatic activity. The recovery of prolylhydroxylase activity during this step was approximately 95%. The contaminating protein, PBP was eluted as a very small peak at an elution volume of about 120 ml (not shown).

3.1.4. Discussion

The purity of the enzyme was assessed by the increase in specific activity, by gel filtration chromatography and polyacrylamide gel electrophoresis. The extent of purification of prolylhydroxylase at different stages of the above described purification scheme is presented in Table 3-1.

As presented in the Table, the initial 0-30% salt precipitation removes as much as 50% of the total proteins present in the original homogenate yet only 17% of the total activity was removed during this stage. The net effect is the concentration of enzymatic activity in the 0-30% supernatant, which accounts for almost 85% of the total activity. Further salt fractionation between 30-60% ammonium sulphate concentration removes about 78% of the proteins originally present in 0-30% supernatant yet only 14% of the activity was lost along with these proteins. The net result is the recovery of 70% of the total activity from the homogenate into the 30-60% pellet and the specific activity is increased by about 7-fold.

In the next stage i.e. poly(Pro)-affinity binding, the recovery of the activity bound to the gel was estimated indirectly by assaying the residual activity left in the supernatant and then by substracting the residual activity from the total activity present in the 30-60% fraction. These studies indicate that as much as

Table 3-1: Purification of Prolylhydroxylase from Chicken Embryos

Step	Fraction	Total Protein, mg	Total Activity U	Activity Recovery %	Specific Activity U/mg	Purification Fold
1	Homogenate	15200	426	100	0.028	1
2	0-30%pellet	8080	73	17	0.01	0.32
3	0-30%supt.	7120	356	84	0.05	1.79
4	30-60%supt.	5615	58	14	0.01	0.37
5	30-60%pellet	1500	298	70	0.20	7.1
6	PP-II	ND	ND	67	ND	ND
7	DEAE	5.4	268	63	50	1773
8	Biogel A-1.5m	3.0	255	60	85	3036

Total protein = total volume x mg protein per ml.

Total activity. = total volume x activity in units/ml

Unit (U) = μmoles of Hyp formed h^{-1}

Specific activity = total activity/total protein

= Units per mg protein

Purification fold = Specific activity at each step/Specific activity at step 1

supt. : supernatant

ND : not determined (due to the presence of bound poly(Pro)

95% of the total activity present in the 30-60% pellet was bound onto the column. This corresponds to about 67% of the total activity present in the homogenate. As mentioned earlier, direct activity measurements on the fractions eluted from the Poly(Pro)-affinity column could not be made since the enzyme is complexed with a potent competitive inhibitor and therefore neither total activity nor specific activity were determined. Hence, the purification fold at this stage remains unknown. The specific activity of the enzyme after the DEAE stage increased tremendously when compared to the 30-60% pellet and the net purification fold at this stage is 1775. The specific activity further rises in the final gel filtration stage increasing the purification fold to 3035. Similar values were reported by Kedersha and Berg (1981). The final specific activity varied from 70-80 units/mg between different batches of the enzyme under the conditions specified in Chapter 2 (See 2.2.2.2). It may be noted that the enzyme lost more than half of its activity within 2 months on several occasions, when it was stored at -20 °C. However, storing the enzyme fractions in liquid N_2 was found to preserve the enzymatic activity for more than 6 months.

3.2. Characterization of Prolylhydroxylase

3.2.1. Molecular Weight Determination

The molecular weight of prolylhydroxylase was determined by calibrating the Biogel A-1.5m column with standard molecular weight markers. The calibration was carried out according to the instructions given in the booklet, supplied along with the molecular weight markers by Pharmacia Chemical Company. The resultant calibration curve is shown in Figure 3-5. As shown in

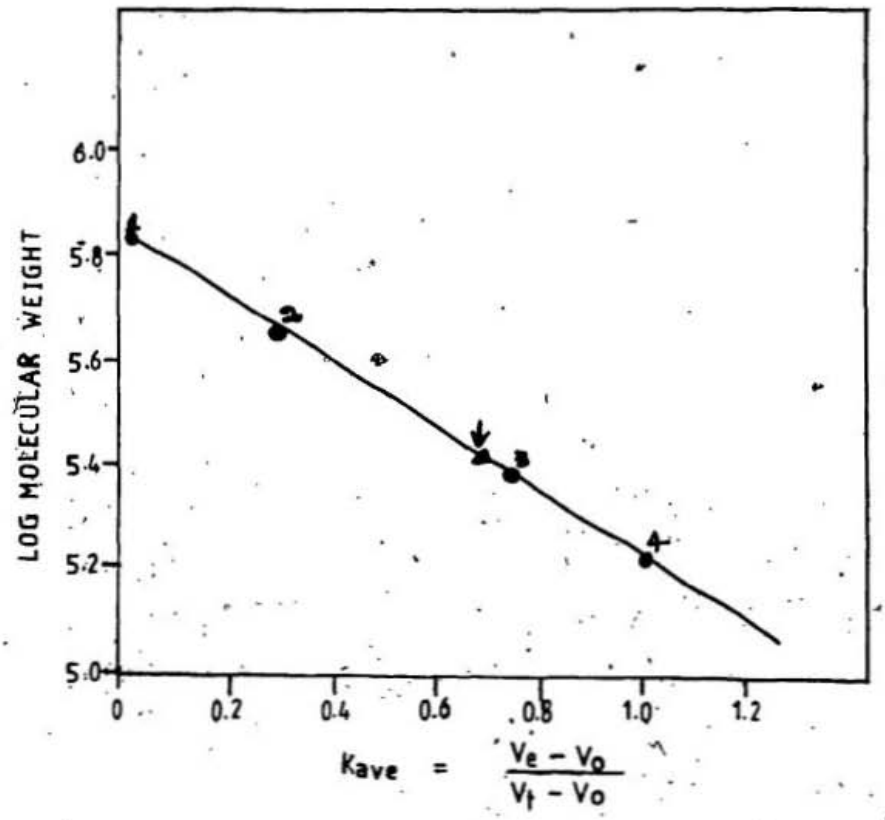


Figure 3-5: Calibration Curve for Biogel A-1.5m Column

The following molecular weight standards were used: (1) Thyroglobulin (M_r 670,000) (2) Ferritin (M_r 440,000) (3) Catalase (M_r 232,000) and (4) Aldolase (M_r 158,000). The closed triangle represents the point corresponding to prolylhydroxylase.

the figure, the K_{av} value of the enzyme is very close to that of catalase and this value translates into a (M_r) of 240,000. Tuderman *et al.* (1975) also reported the same value for the chicken enzyme from amino acid composition and electrophoresis under non-denaturing conditions. In this context, it may be noted that human prolylhydroxylase also has the same M_r of 240,000 (Kuutti *et al.*, 1975).

3.2.2. Subunit Structure

As discussed in Chapter 1 (section 1.6.1), prolylhydroxylase was reported to be a tetramer made up of two α-subunits and two β-subunits with M_r of 64,000 and 61,000 respectively. The subunit structure and the M_r of these subunits was confirmed by SDS-PAGE on slab gels. The photograph of the gel is presented in Figure 3-6 and the corresponding calibration curve in Figure 3-7.

In Figure 3-6, lane-A represents the enzyme fraction purified on the DEAE ion-exchange column. This clearly shows the band due to the contamination from PBP, in addition to the bands corresponding to α- and β subunits. Lane B and D represent the molecular weight markers in the order of decreasing molecular weights. Lane C corresponds to the final enzyme fraction purified by gel filtration chromatography on Biogel A-1.5m column and clearly demonstrates the removal of contaminating PBP. One can also see from the photograph that purified chicken prolylhydroxylase is composed of equal amounts of α and β subunits.

As can be seen from the calibration curve (Figure 3-7), the R_f values of the enzyme subunits correspond to monomer M_r values of 64,000 and 60,000. These

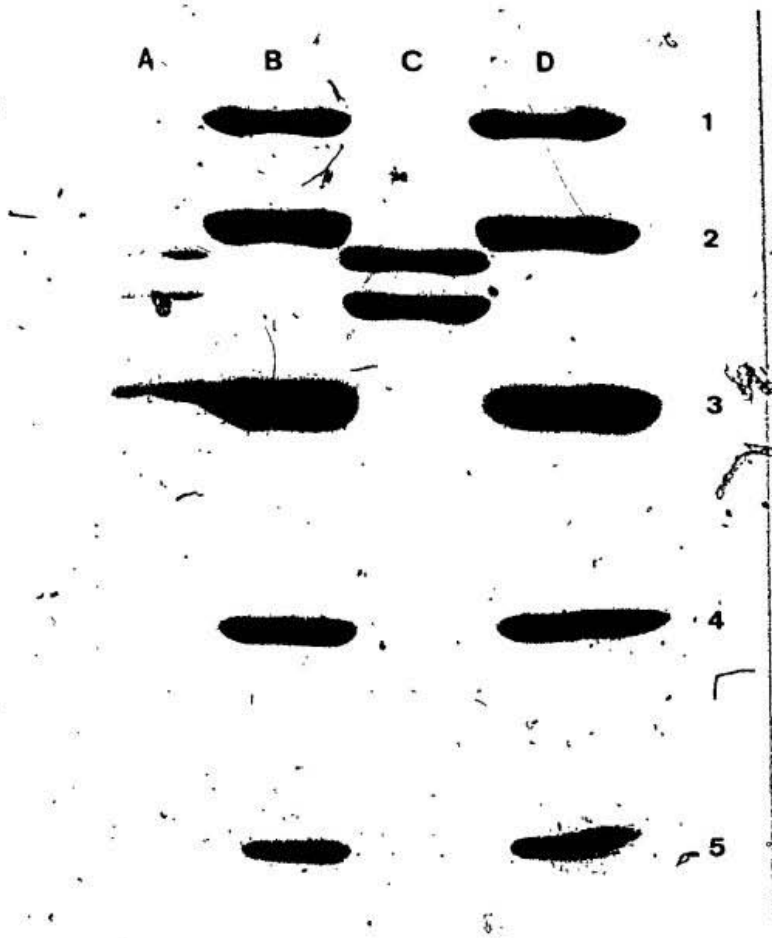


Figure 3-6: SDS-PAGE showing the Subunit Composition of Prolylhydroxylase

Lane A contains enzyme fraction purified from DEAE column;

Lane C contains final enzyme fraction obtained from Biogel column.

Lane B and D contain molecular weight markers. (1) Phosphorylase (M_r 94,000) (2) Bovine serum albumin (M_r 67,000) (3) Ovalbumin (M_r 43,000 (4) Carbonic anhydrase (M_r 30,000) and (5) α-Lactalbumin (M_r 14,400)

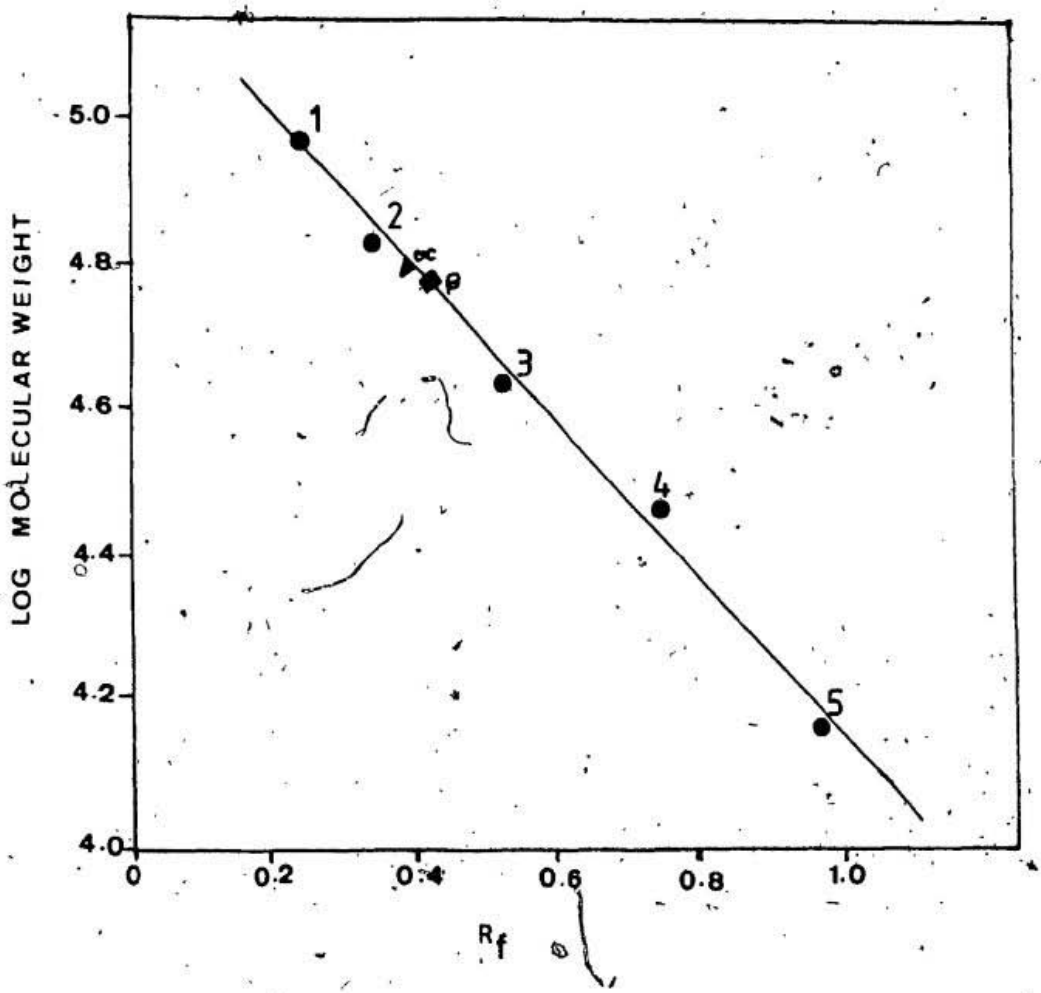


Figure 3-7: Calibration Curve for SDS-PAGE

The molecular weights of the standards are as given in the legend for Figure 3-6. The closed triangle represents that of α- subunit and the closed square represents β-subunit.

values are very close to those published earlier by Kuutti *et al.* (1975) and Berg *et al.* (1979) for human and chicken enzyme.

3.3. Standardization of Prolylhydroxylase Reaction

The reaction conditions for the prolylhydroxylase reaction were standardized by studying the effects of the various factors that influence the initial velocity of the enzymatic activity and the optimal conditions for enzymatic assay were established.

3.3.1. Effect of Time

Figure 3-8 shows the plot of prolylhydroxylase activity as a function of time. As can be seen from the figure, the enzymatic activity was found to be linear until 20-25 min and the standard assay was carried out routinely for 15 min. However, for the sake of uniformity in comparing the results of different experiments, in what follows, v is expressed as the μmol of Hyp produced in 60 min (1 h) by 1 mg of enzyme (i.e μmol/h/mg); v can therefore, also be referred to as the specific activity.

3.3.2. Effect of Temperature

Figure 3-9 shows the plot of prolylhydroxylase activity as the function of temperature. As seen from the figure, using the assay conditions mentioned in the legend, the optimum temperature for the maximum activity is about 37 °C.

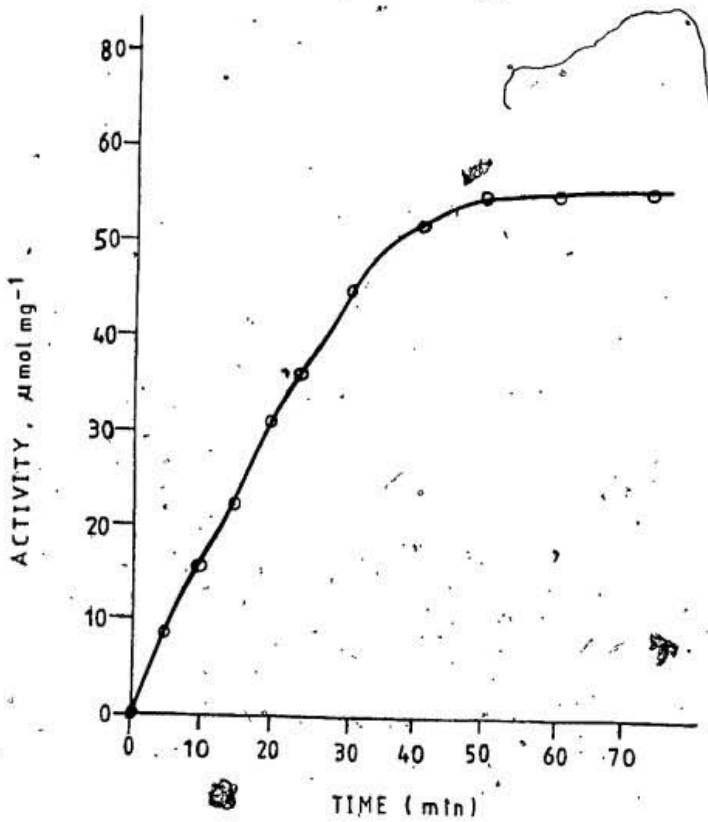


Figure 3-8: Effect of Time on Prolylhydroxylase Activity

The reaction was carried out under the conditions described in the legend to Figure 3-9 for the time intervals shown above at 37 °C.

Each point represents the average of atleast 3-4 individual trials.

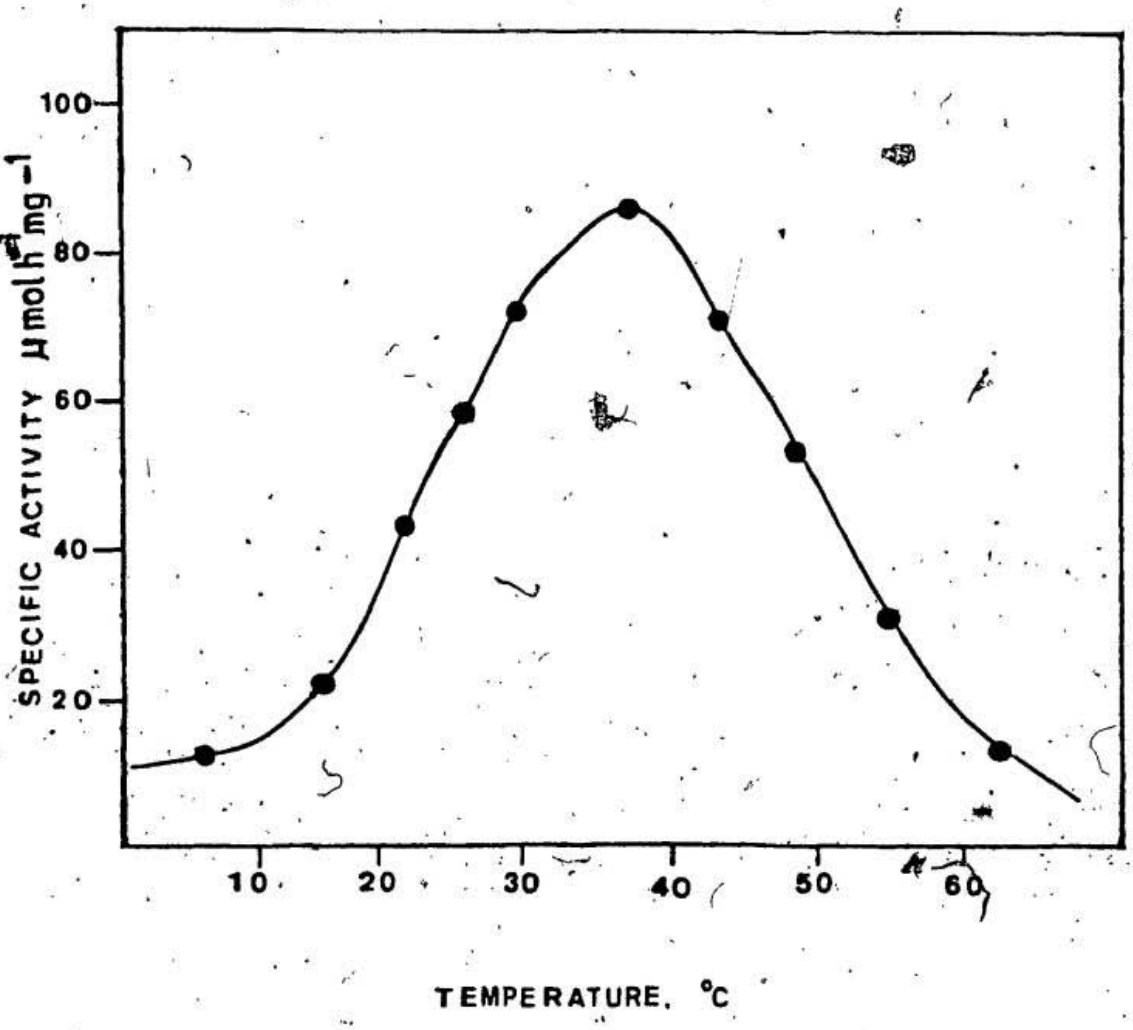


Figure 3-9: Effect of Temperature on Prolylhydroxylase Activity

The assay of prolylhydroxylase was carried out at different fixed temperatures. The 1 ml assay medium contains 50 mM Tris-HCl pH 7.4; 2 mg/ml BSA, 0.1 mg/ml catalase, 0.1 mM DTT, 2 mM ascorbate, 5-6 μg enzyme, 0.1 mM $FeSO_4$, 2 mM α-KG and 0.5 mg/ml $(Pro\text{-}Pro\text{-}Gly)_n$. Reaction time is 15 min; Each point represents the average of atleast 3-4 individual trials.

3.3.3. Effect of pH

The pH optimum for the prolylhydroxylase reaction was investigated by carrying out the assay at different pH values. Preparations of 50 mM Tris-HCl buffer of different pH values were prepared at 25 °C. In view of the known pH variation of Tris buffers with temperature, the pH values of the buffer solutions were measured again at 37 °C. Standard reaction mixtures were prepared in 50 mM Tris-HCl of different pH values measured at 37 °C and the reaction was carried out as usual at 37 °C. Figure 3-10 shows the plot of prolylhydroxylase activity as a function of pH. From these results, it is found that the pH optimum for prolylhydroxylase at 37 °C is about 7.3-7.4. This value is in excellent agreement with the value reported by Kivirikko and Prockop (1967a) for chicken enzyme.

3.3.4. Effect of Enzyme Concentration

The effect of increasing amounts of enzyme on the initial velocity of the reaction was investigated and the results are presented in Figure 3-11. As seen from the figure, the initial velocity of the reaction increases linearly with increasing amounts of prolylhydroxylase until a concentration of about 25 μg/ml. In the standard assay however, about 5-10 μg per ml of the reaction mixture was used.

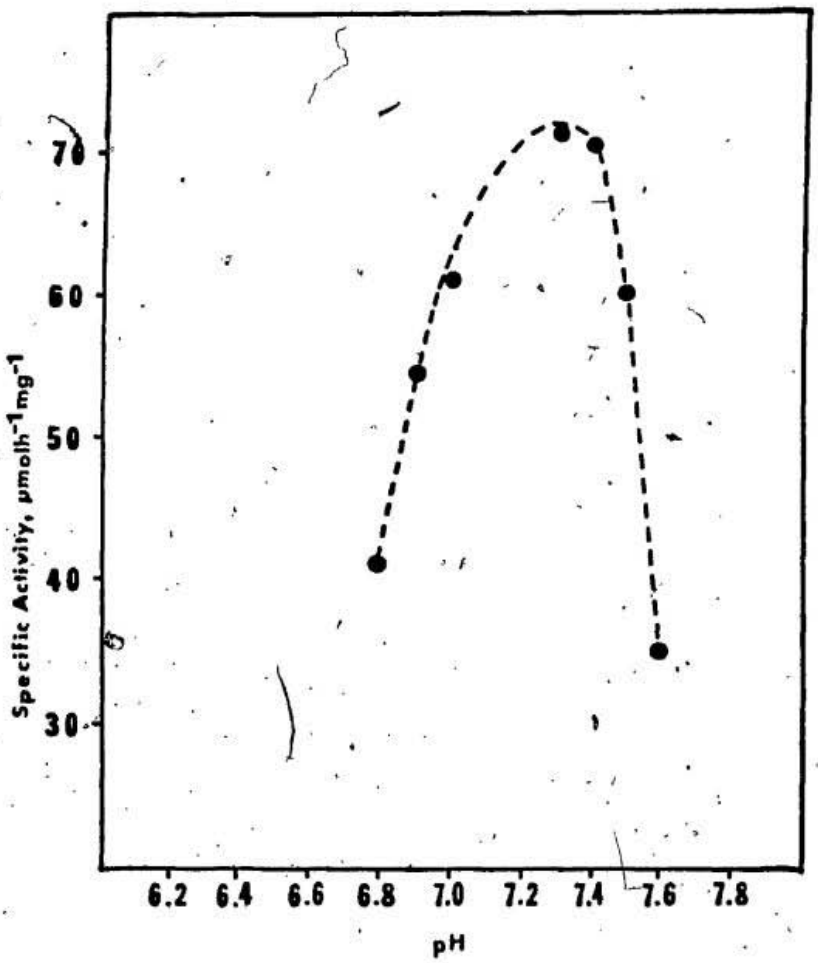


Figure 3-10: Effect of pH on Prolylhydroxylase Activity

The assay of prolylhydroxylase was carried out at different pH values of the reaction medium. The temperature was fixed at 37 °C. The assay conditions were as described in the legend to Figure 3-9. Each point represents the average of atleast 3 individual trials.

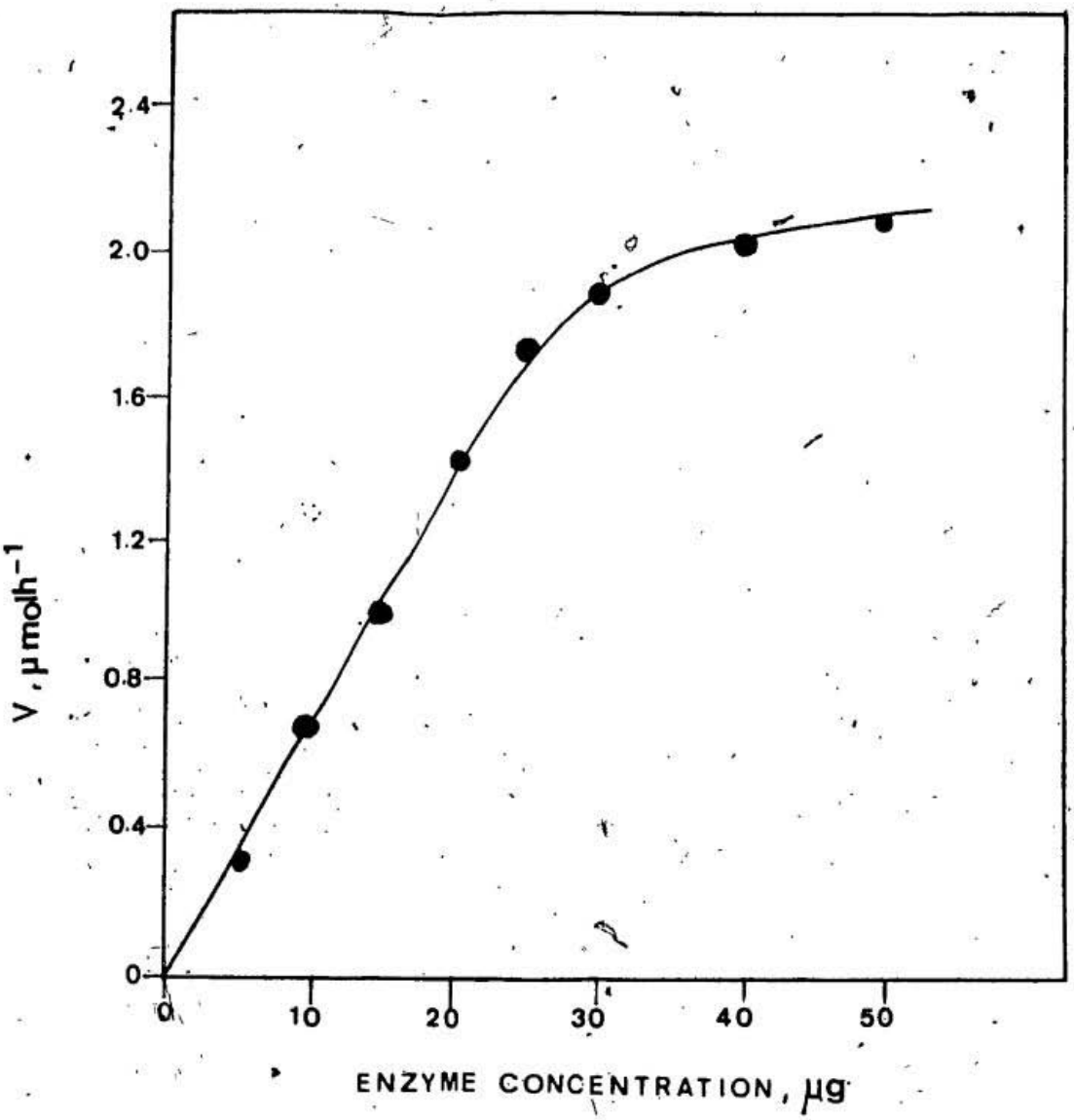


Figure 3-11: Effect of Enzyme Concentration on the Initial Velocity of Prolylhydroxylase Reaction

The reaction conditions were as described in the legend for Figure 3-9. The concentration of the enzyme was varied as above and the temperature was fixed at 37 °C. Each point represents the average of atleast 3 individual trials.

3.3.5. Effect of Substrate Concentration

The effect of increasing amounts of the substrate $(Pro\text{-}Pro\text{-}Gly)_5$ was investigated in the concentration range of 0-2 mM. The results are presented in Figure 3-12. As seen from the Figure, the plot of initial velocity versus substrate concentration displays a near-linear initial part followed by saturation at substrate concentrations i.e. beyond 1 mM. Figure 3-13 shows the double reciprocal plot of the same data. From this, a K_m value of 0.5 mM and the maximum velocity (V_{max}) of 111 μmol/h/mg, at infinitely high substrate concentration were obtained. Similar values of K_m (0.55$\pm$0.075 mM) and V_{max} (120$\pm$15 μmol/h/mg) were obtained on a number of occasions using different batches of enzyme preparations. Similar values have been reported by Nietfield and Kemp (1980). However, Myllyla *et al.* (1977) and Berg *et al.* (1977) reported somewhat lower K_m values of 0.2-0.3 mM, although V_{max} values were similar to the ones obtained here. This apparent discrepancy may be due to the different cosubstrate and cofactor concentrations used in the reaction mixtures by these authors when compared to the conditions employed here (see section 3.4). According to Engel (1981), in multi-substrate enzyme mechanisms, the apparent K_m for one substrate depends on the concentration of the other substrate and this may well account for the differences observed for K_m values reported.

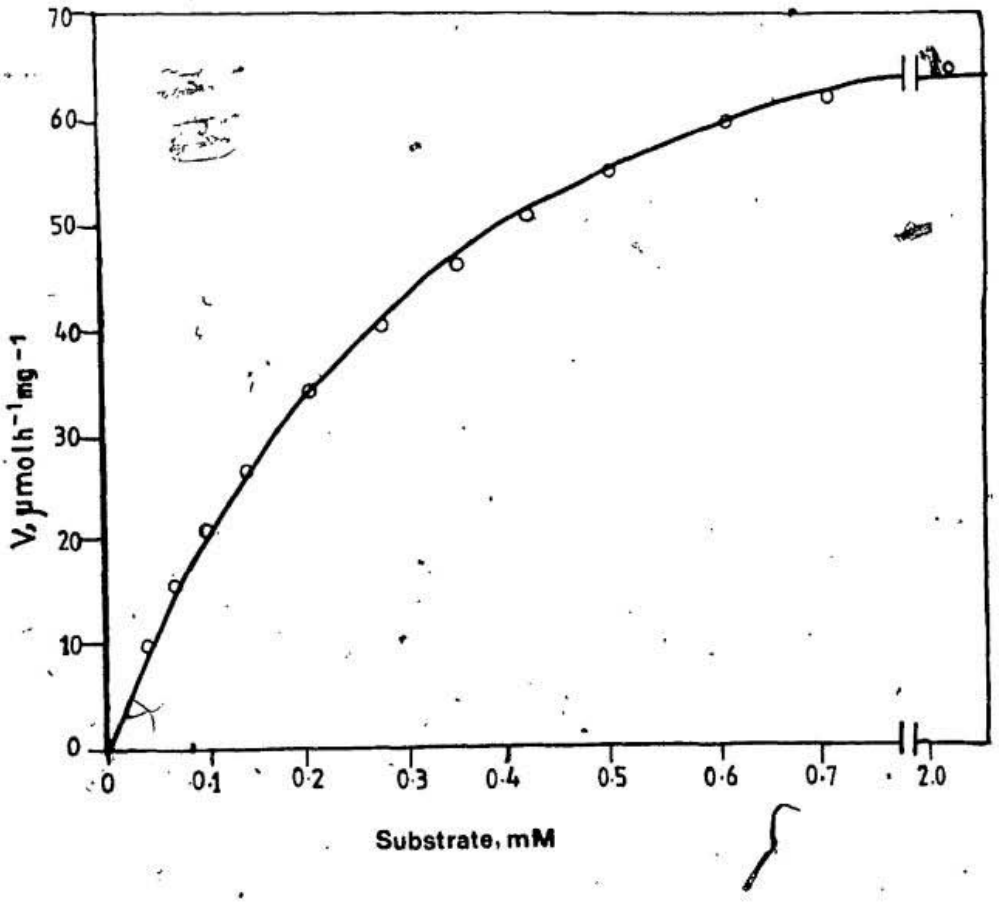


Figure 3-12: Effect of Substrate Concentration on the Initial Velocity of Prolylhydroxylase Reaction

The reaction was carried out as described in the legend for Figure 3-9 except that substrate concentration was varied as shown above. Temperature was fixed at 37 °C. Each point represents the average of atleast 5 individual trials.

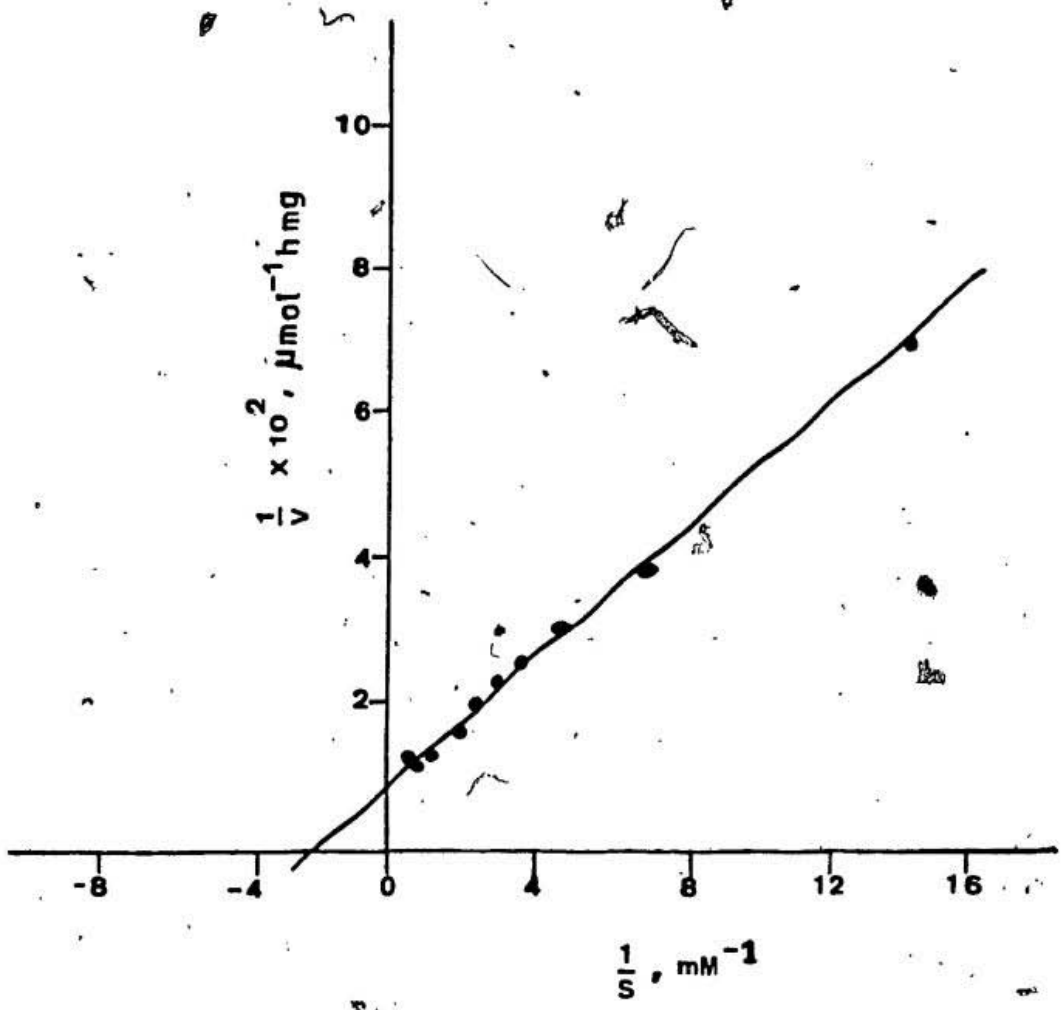


Figure 3-13: Lineweaver-Burk Plot for the Effect of Substrate Concentration on Prolylhydroxylase Reaction

Line is drawn according to linear regression analysis with correlation coefficient of 0.991. Each point represents the average of atleast 5 individual trials.

3.3.6. Effect of α-KG Concentration on the Initial Velocity

The effect of increasing amounts (0-5 mM) of the cosubstrate, namely, α-KG on the initial velocity of prolylhydroxylase reaction was studied. The results are presented in Figure 3-14 and Figure 3-15.

As seen from the Figure 3-14, the reaction velocity shows linear response with increasing concentrations of α-KG up to 1 mM. The maximal velocity was obtained between 1-2 mM beyond which, the reaction starts decreasing. At the maximal α-KG concentration studied (5 mM), the activity was decreased by about 30% of the maximal value.

The double reciprocal plot of this data is shown in Figure 3-15 which clearly shows the inhibition of the reaction at high cosubstrate concentration deviating from both linearity as well as saturation. The K_m value obtained from this data is 0.15 mM and the K_i value at the high cosubstrate concentration is about 2.5 mM. Inhibition at high cosubstrate concentration has not been reported earlier.

3.3.7. Effect of $FeSO_4$ Concentration on the Initial Velocity

The effect of increasing amounts of the cofactor, $FeSO_4$ (0-0.2 mM) on the initial velocity of prolylhydroxylase reaction was studied in the presence of two different concentrations of α-KG i.e. 0.1 mM and 2.0 mM which is the saturating concentration as observed in the earlier section (Figure 3-14). The results are presented in Figures 3-16 and 3-17. As seen from Figure 3-16, in the presence of 0.1 mM α-KG, the activity increases linearly with increasing concentrations of $FeSO_4$ and the maximum activity was obtained between 0.04-0.05 mM of $FeSO_4$.

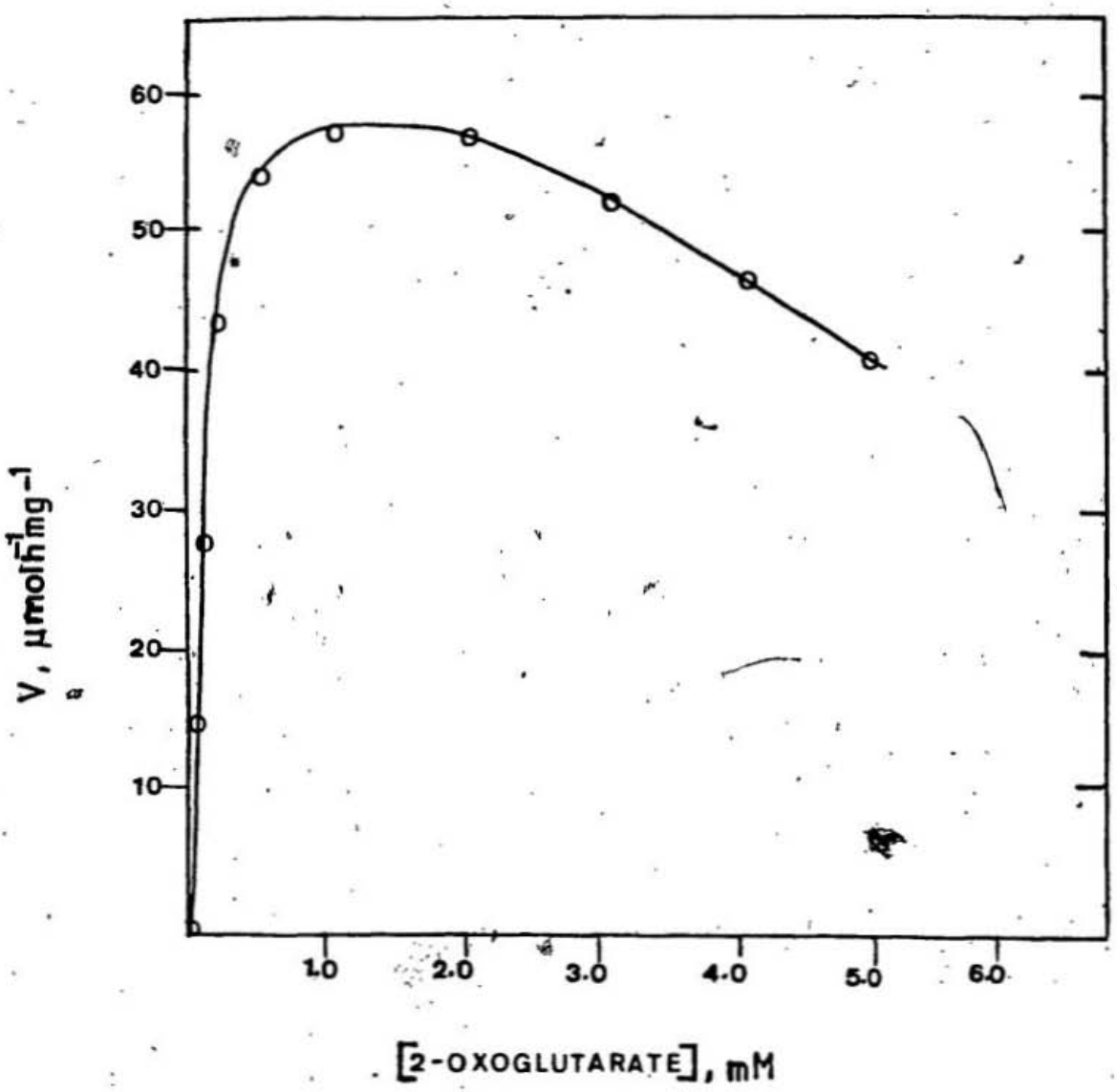


Figure 3-14: Effect of α-KG Concentration on the Initial Velocity of Prolylhydroxylase Reaction

The assay was carried out under the conditions described in the legend for Figure 3-9 except that the concentration of α-KG was varied as shown above; temperature was fixed at 37 °C; each point represents the average of atleast 3 individual trials.

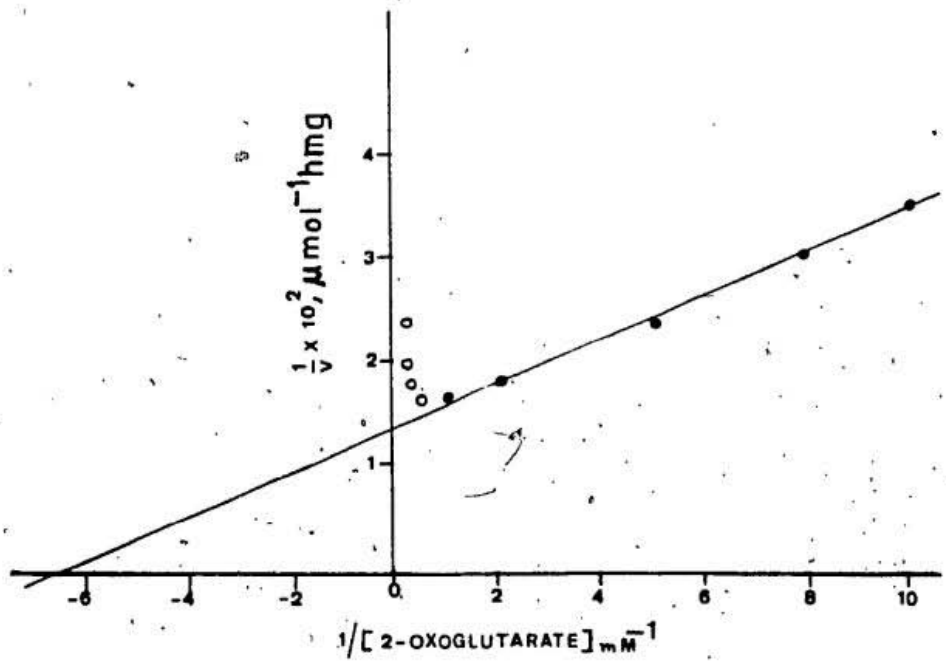


Figure 3-15: Double Reciprocal Plot for the Effect of α-KG Concentration on Prolylhydroxylase Reaction

Open circles are the points in the high cosubstrate range showing inhibition at these concentrations and the closed circles are in the lower cosubstrate concentration range. Line is drawn according to the linear regression analysis with a correlation coefficient of 0.985 for the points in the lower cosubstrate concentration range.

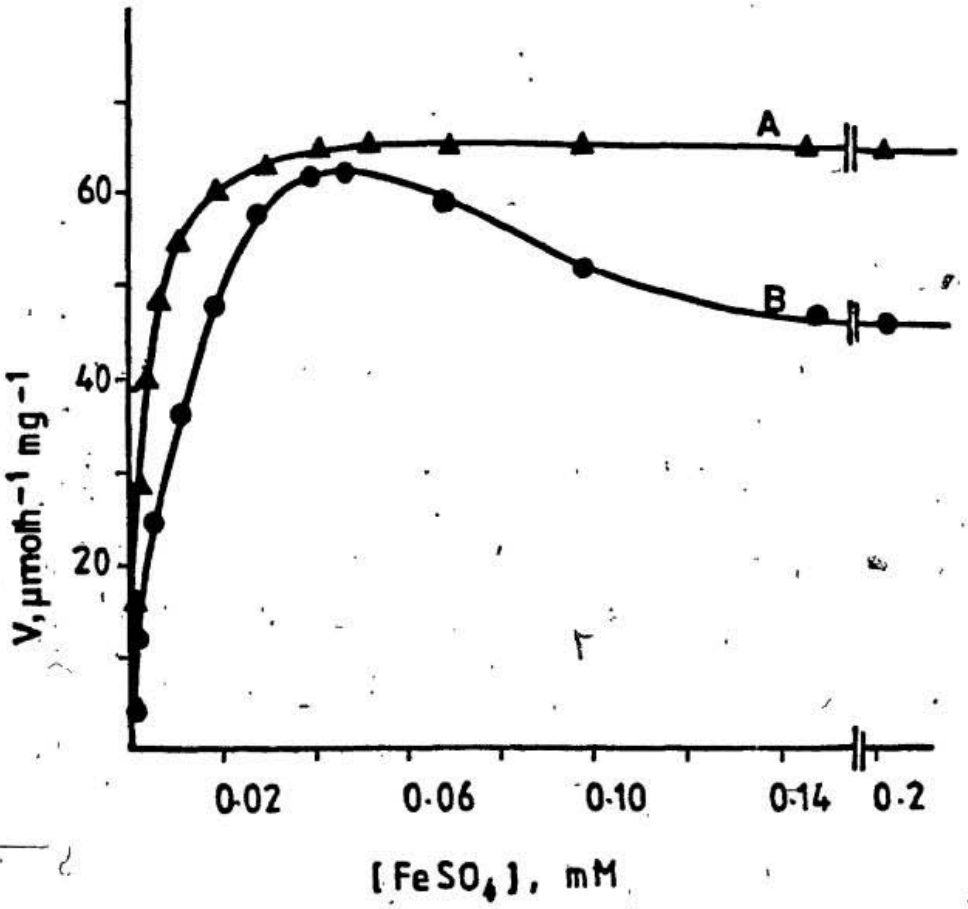


Figure 3-16: Effect of $FeSO_4$ on the Initial Velocity of Prolylhydroxylase Reaction

The assays were carried out as described in the legend of Figure 3-9 except that the concentration of $FeSO_4$ was varied as shown above. Two separate experiments were done; the $FeSO_4$ concentration was varied with fixed concentrations of α-KG at A. 2 mM (line with closed triangles) and B. 0.1 mM (line with closed circles).

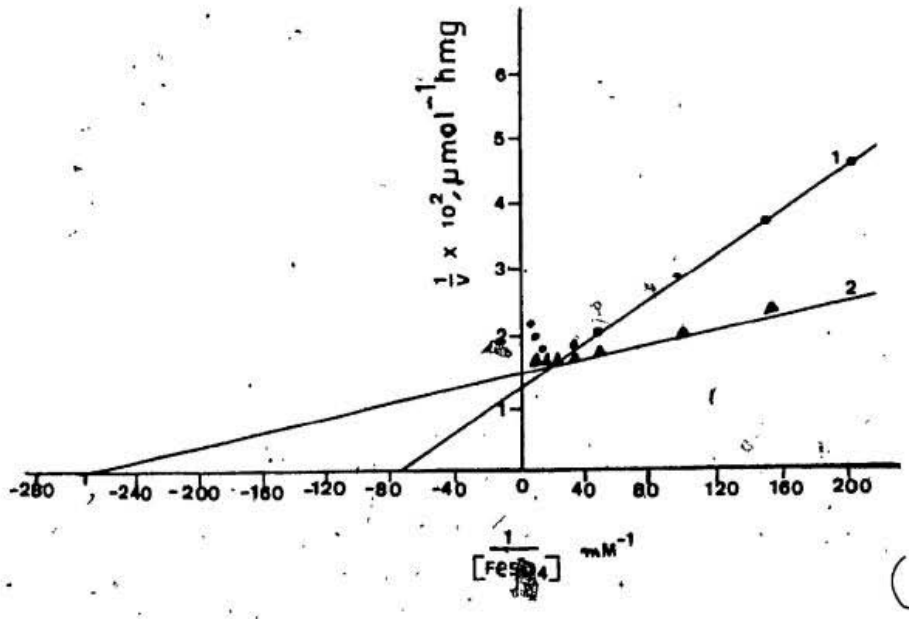


Figure 3-17: Double Reciprocal Plots for the Effect of $FeSO_4$ on Prolylhydroxylase Reaction

Assays were carried out as described earlier; Line 1 represents the effect of increasing amounts of $FeSO_4$ in the presence of 0.1 mM α-KG (closed circles) and line 2 represents the effect of $FeSO_4$ concentration in the presence of 2.0 mM α-KG (closed triangles), on the initial velocity of prolylhydroxylase reaction. Lines were drawn according to linear regression analyses. In the case of closed triangles, all the points were included while with closed circles, points beyond 0.05 mM $FeSO_4$ were eliminated.

Beyond 0.05 mM, the activity gradually decreases to a plateau at about 0.2 mM $FeSO_4$ concentration. The activity at this point is only about 70% of the maximal value obtained at 0.04-0.05 mM. A similar behaviour was reported by Tuderman *et al.* (1977a).

When a saturating concentration (2 mM) of α-KG was used, the velocity quickly reached the maximum at about 0.04-0.05 mM $FeSO_4$. However, unlike the earlier case, the velocity was maintained at the maximum throughout the further concentration range between 0.05--0.2 mM. These results indicate that in the presence of 0.1 mM α-KG, although the maximum velocity was obtained at about 0.05 mM $FeSO_4$, the velocity soon decreases probably because the concentration of α-KG becomes rate-limiting. On the other hand, when saturating concentrations of α-KG are present, the maximal velocity is sustained. When these data were plotted as double reciprocal plots (Figure 3-17) one can clearly see the effect of different concentrations of α-KG, on the apparent K_m values of enzyme for $FeSO_4$. With 0.1 mM α-KG, a value of 0.0138 mM (13.8 μM) was obtained while with saturating concentrations of α-KG, the K_m value for $FeSO_4$ decreased 4-fold resulting in 3.8 μM. The latter value is very close to that (5 μM) reported earlier by other investigators (Tuderman *et al.*, 1977a; Nixtfield and Kemp, 1980).

One can also see from Figure 3-16, that the presence of exogenously added $FeSO_4$ is not an absolute requirement for the enzymatic proline hydroxylation. There seems to be some activity by the enzyme (about 6.5% with 0.1 mM α-KG and as much as 40% with 2 mM α-KG) in the absence of added $FeSO_4$. This

activity can be considered endogenous due to the bound ferrous ions of the native enzyme. Since the water and buffer solutions were deionized by passing through the metal-ion retarding resin, Chelex-100, the possibility of minor contamination by extraneous ferrous ions is ruled out.

3.4. Summary and Discussion

In order to carry out the proposed studies on the interaction of prolylhydroxylase with peptides (Chapter 6) and on the characterization of the enzyme's structure (Chapter 4), it was necessary to obtain the enzyme in the purest possible form and to assess its kinetic and other properties. Literature reports on the purification and characterization (especially kinetic) are not always consistent. This has resulted in problems of comparing the data on peptide-enzyme interactions reported from different laboratories. Therefore, considerable care was taken here to obtain consistently pure enzyme preparations with high specific activity and well defined kinetic parameters.

Prolylhydroxylase has been purified from 13-14 day old chicken embryos using an affinity procedure after salt fractionation at 30-60% ammonium sulphate concentration. The enzyme was further purified by DEAE ion-exchange chromatography and gel filtration on Biogel A-1.5m. The enzyme obtained by this procedure was about 3,000-fold pure and was found to be homogenous with a high specific activity of 70-80 units in terms of μmol Hyp produced by 1 mg of enzyme in one hour under the conditions specified in section 2.2.3.2. However, under saturating conditions of the substrate, the maximal specific activity varied around 120 (±15) μmol/h/mg (or 1.8-2.3 μmol/min/mg) of the enzyme, which is

very similar to the values reported by other investigators (Tuderman *et al.*, 1977a; Nietfield and Kemp, 1980 and Kedersha and Berg, 1981). If this value is used for the calculation of enzyme activity as was done by Kedersha and Berg (1981), the purification fold further increases by as much as 1,000 fold resulting in the net purity of about 4,500-fold. This value is close to the value reported (5,000-fold) by the same authors.

The enzyme exhibited a M_r of about 240,000 by gel filtration and has a subunit composition of α_2 β_2. The α-subunits exhibit a M_r of 64,000 and β-subunits of 60,000 as determined from SDS-PAGE.

Prolylhydroxylase activity is rather sensitive to the changes in temperature and pH of the reaction medium. The optimum temperature was found to be around 37 °C and the optimum pH was about 7.3-7.4 at this temperature. These effects may be related to the effects of temperature a and pH on the structural stability of the enzyme protein as will be discussed in Chapter 6. The pH optimum suggests the involvement of His residues in the enzyme's active site, although this speculation needs further confirmation from other direct studies.

The initial velocity was found to be linear up to 20-25 min under the standard conditions specified in Chapter 2. Increasing amounts of enzyme up to 25 μg/ml produced a linear response under the conditions of substrate and cofactors used and beyond 25 μg/ml, the rate assumed a nearly zero-order kinetics.

Prolylhydroxylase is completely dependent on the α-KG concentration for

activity and no activity was observed in the absence of added α-KG. Maximal velocity was obtained between 1-2 mM of α-KG and higher concentrations of this cosubstrate seemed to inhibit the reaction. The K_m value obtained was 0.15 mM for α-KG. This value is one order of magnitude higher than those values reported by other investigators (0.01-0.02 mM) (Tuderman *et al.*, 1977a). The reason for this discrepancy is not clear. Inhibition at higher cosubstrate concentrations has not been reported earlier. Probably, high concentrations of succinate generated from α-KG may be responsible for this inhibition (end-product inhibition).

Prolylhydroxylase, as purified, was capable of carrying out hydroxylation reaction to some degree in the absence of added ferrous iron. While Tuderman *et al.* (1977a) reported the absolute requirement for exogenous ferrous ions, Nietfield and Kemp (1980) reported that their preparation of prolylhydroxylase was able to carry out hydroxylation in the absence of added iron. The latter authors showed that prolylhydroxylase purified by the affinity procedure still contains firmly bound iron that is responsible for the observed prolylhydroxylase activity, in the absence of added $FeSO_4$ (Nietfield and Kemp, 1980). Later studies by De Jong and Kemp (1982) reported that at maximal activity (2.0 μmol/min/mg), the enzyme contains two Fe^{+2} ions specifically bound per mole of enzyme tetramer. They also showed that the K_m value for $FeSO_4$ depends on the nature of the sulphhydryl compounds present in the reaction medium. Cysteine was found to stimulate the enzymatic activity in the absence of added iron and BSA (De jong and Kemp, 1982). These authors also reported that in the absence of added $FeSO_4$, BSA and DTT could stimulate the enzymatic activity by 5-10% of the maximal rate. A relatively high rate, as much as 50% of the maximal activity,

was reported by these authors in the absence of added Fe^{+2} if BSA was preincubated with DTT. The stimulation by cysteine or DTT suggests that the free SH groups are essential for the enzymatic activity. This suggestion is supported by the observation that the enzyme is inhibited by low concentrations of thiol reagents (Popenoe *et al.*, 1969; Halme *et al.*, 1970), an inhibition that can be prevented by α-KG (Popenoe *et al.*, 1969) or reversed by DTT (Halme *et al.*, 1970). According to Kivirikko and Myllyla (1980), the action of BSA is in part explained by a nonspecific "protein effect", but in part, more specifically due to the presence of a number of free SH groups on BSA. These considerations may also offer an explanation for the observed activity in our studies in the absence of added Fe^{+2} because in the reaction medium, BSA and DTT are added in this order before the addition of the enzyme. Therefore, the enzyme may have been stimulated, in the absence of Fe^{+2}, by BSA treated with DTT. The reaction medium also contains catalase which was found to stimulate the prolylhydroxylase reaction. The effect of catalase is partly due to the destruction of peroxide, which is generated non-enzymatically by solutions of O_2, $FeSO_4$ and ascorbate (Kivirikko and Prockop, 1967c).

The K_m value for $FeSO_4$ was about 14 μM when 0.1 mM α-KG was used while with saturating concentrations of α-KG (2 mM), the K_m value decreased by 4-fold resulting in a value of about 4 μM. Although similar activities were obtained either with 0.1 or 2.0 mM α-KG at 0.05 mM concentration of $FeSO_4$, the activity began to decrease beyond 0.05 mM $FeSO_4$ in the presence of 0.1 mM α-KG probably because the concentration of the latter becomes rate limiting.

incubation for 3-5 min at 0-4 °C, 0.1 mg/ml catalase, 2 mM ascorbate, prolylhydroxylase (5-10 μg/ml), 0.1 mM $FeSO_4$ and, after at least 15-30 sec, 2 mM α-KG. The reaction mixture was then transferred from the ice-bath to 37.5 °C in a shaking water bath and preequilibrated for 8 min; finally 0.35 mM $(Pro\text{-}Pro\text{-}Gly)_5$ was added to start the reaction. The reaction time was 15 min during which the reaction remains linear. These conditions were found to give optimal enzymatic activities.

Chapter 4

Conformational Aspects of Prolylhydroxylase

4.1. Structural Properties of Prolylhydroxylase

Although prolylhydroxylase has been purified to homogeneity over the past several years, there have been very few structural studies on this enzyme. Prolylhydroxylase from chicken embryos, new born rat and human sources has been shown to be a tetramer consisting of two different types of enzymatically inactive subunits with M_r of 64,000 and 60,000, respectively (Tuderman *et al.*, 1975; Kuutti *et al.*, 1975; Kedersha and Berg, 1981). At the time of the begining of these structural studies on the pure enzyme (presented in this Chapter), the only structural study available was the electron microscopic study carried out by Olsen *et al.* (1973) which showed that the monomers α and β are rod-shaped and are joined to form V-shaped dimers which are interlocked to form tetramers of $\alpha_2\beta_2$. Not much information was available from chemical modification studies either, except for the fact that the disulphide bonds are essential for maintaining the native structure and activity of the enzyme (Berg *et al.*, 1979).

Considering the fact that collagen is the major protein in vertebrates amounting to over 20% by weight of all proteins, the paucity of data on the

structure of one of the key enzymes in the biosynthesis of this protein is indeed surprising. Since the structure is related to function, it should be interesting to study this aspect of prolylhydroxylase in the light of the high degree of specificity it exhibits towards the conformation of its polypeptide substrates, as discussed in the introductory chapter. As part of the attempts to understand the interaction of prolylhydroxylase with the peptide substrate, cosubstrate (α-KG) and cofactor ($FeSO_4$), the secondary and the tertiary structural characteristics of prolylhydroxylase have been determined using spectroscopic techniques (CD and fluorescence spectroscopy) in the absence and presence of these reactants. The results obtained are presented in this Chapter.

4.1.1. Conformation of Prolylhydroxylase from CD Studies

CD spectra of proteins in the wavelength regions of 180-250 nm (far-UV spectra) and 250-320 nm (near-UV spectra) give information, respectively, about the secondary and tertiary structural characteristics. While the secondary structural features are due to the conformation of the backbone, the tertiary structural features are due to the relative orientation of these secondary structural elements in space and thereby indicative of the internal environment of the protein molecule (Schulz and Schirmer, 1979). Therefore, in the present studies, CD measurements of prolylhydroxylase were used for characterizing the conformational features of prolylhydroxylase. The far- and near-UV CD spectra of the enzyme were obtained using a Jasco J-500 A spectropolarimeter equipped with a DP-500 N data processor (Section 2.2.4.3) and the results are presented below.

The far-UV CD spectrum of prolylhydroxylase at 0-2 °C and pH 7.8 is shown in Figure 4-1. The presence of two minima at 208 and 220 nm respectively, and a maximum at about 192 nm characterizes the spectrum. The spectrum shown represents the average of several spectra run with different batches of the enzyme having maximal specific activities i.e. in the range of 90-110 units/mg. The ellipticity values are -16,000 $\pm$ 1,000 deg.cm^{2}dmol^{-1} at the minima and 19,000 $\pm$ 1,000 deg.cm^{2}dmol^{-1} at the maximum, as obtained with different batches of the enzyme preparations. These values were found to be practically invariant between 0-20°C and hence could be taken to represent the native enzyme. These spectral features are indicative of significant amounts of α-helical conformation in the native enzyme.

The α-helical content was calculated from the observed ellipticity of the enzyme at 208 nm and from the ellipticity values, at this wavelength, of the random-coil and fully α-helical conformations of poly-L-lysine in water (Greenfield and Fasman, 1969). An average value of 38 $\pm$ 2% was observed for the α-helical content of native prolylhydroxylase between 0-20 °C. Comparison of the far-UV CD spectrum of prolylhydroxylase with those computed by Greenfield and Fasman (1969) for various mixtures of the α-helix, β-sheet and random-coil conformations of the model polypeptide, namely, poly-L-lysine, indicated that the far-UV CD spectrum representing about 40% α-helix, 40% β-sheet and 20% random-coil is very similar to that obtained for native prolylhydroxylase. Based on these observations, one can conclude that prolylhydroxylase has a substantial proportion of α-helical conformation (about 40%). More detailed analysis of the

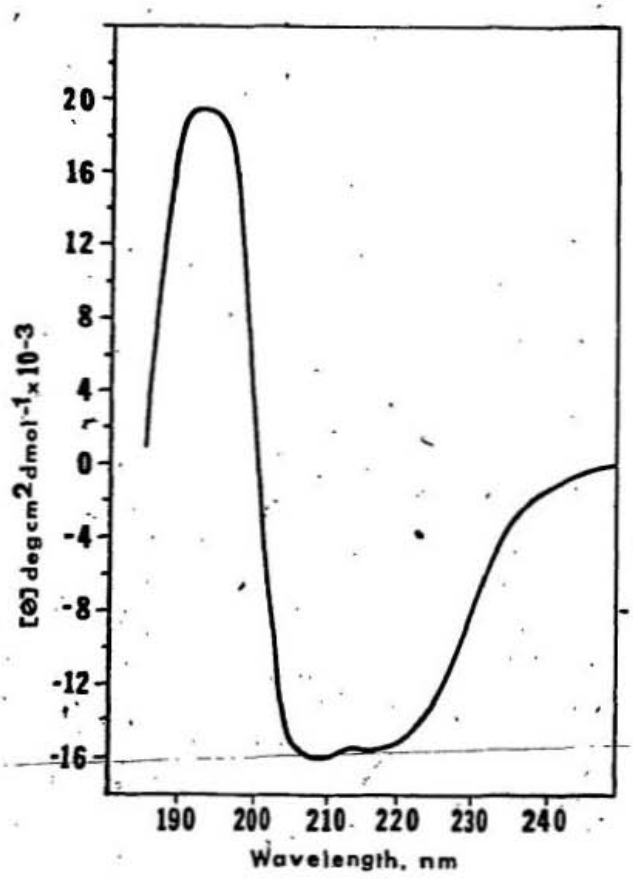


Figure 4-1: Far-UV CD Spectrum of Prolylhydroxylase

spectrum was obtained at 0-2 °C in Tris-buffer, pH 7.8.

Concentration = 0.5 mg/ml; cell pathlength = 0.1 cm.

conformational composition from CD spectra involves the use of the procedures based on the CD spectral data of globular proteins of known secondary structures from X-ray crystallography. This was attempted using the procedure of Prof. Martinez (Chen *et al.*, 1972) with the help of Dr. Devarajan at the University of California, San Fransisco. The results obtained are 42% α-helix, 0% β-sheet, 16% β-turns and 42% random-coil structure. It is very interesting to note that both the above methods give identical values for the proportion of α-helix present in prolylhydroxylase.

The near-UV CD spectrum of prolylhydroxylase at 0-2 °C and pH 7.8 is shown in Figure 4-2. A series of negative CD bands are observed between 250-320 nm region which can be attributed to asymmetrically oriented aromatic amino acid residues (Adler *et al.*, 1973). Based on the available data (Adler *et al.*, 1973), those around 262 nm and 268 nm may be attributed to the Phe residues while the bands at 275 and 282 nm may be assigned to tyrosyl residues placed in an asymmetric environment in native prolylhydroxylase. The negative band at 288 nm could arise either from tyrosyl, or more likely, from tryptophanyl residues. The amino acid compostion data on the enzyme (Berg *et al.*, 1979) show that prolylhydroxylase contains as many as 68 tyrosyl and 130 phenylalanyl residues (per M_r of 240,000). No data are, however, available on the number of tryptophanyl residues, if any, present in the enzyme. It would appear that this is due to the special procedures needed for estimating the tryptophanyl residue content. Therefore, the absence of such data does not necessarily indicate the absence of tryptophanyl residues. As will be shown in a subsequent section, a definitive indication of the presence of tryptophanyl residues in the enzyme is

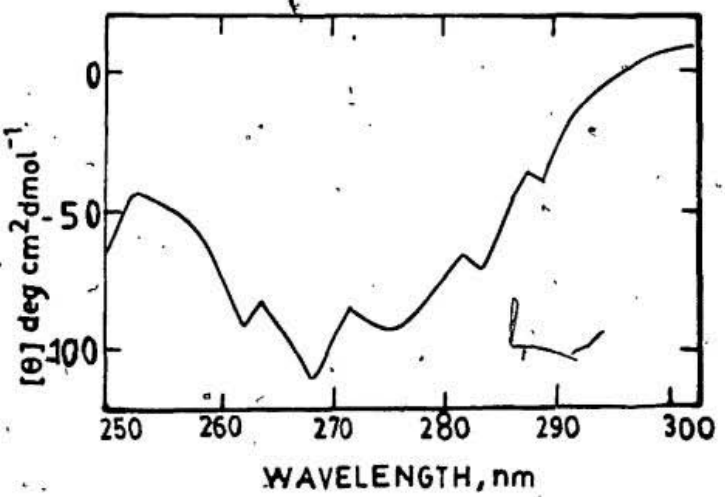


Figure 4-2: Near-UV CD Spectrum of Prolylhydroxylase

spectrum was obtained at 0-2 °C in Tris-buffer, pH 7.8.

Concentration = 0.5 mg/ml, cell pathlength = 1 cm.

observed from the fluorescence data. More studies involving, for example, specific chemical modifications are necessary, however, to interpret the 288 nm CD band in terms of tryptophanyl contribution. In any case, the aromatic CD bands of prolylhydroxylase should be useful in studying its interactions with substrates, inhibitors and cofactors in terms of the perturbation of the environment of the aromatic residues.

4.1.2. Structural Data from Fluorescence Studies

CD measurements monitor the optical activity of the proteins due to the conformation of the peptide back-bone and to molecular asymmetry (Bayley, 1980). Both secondary and tertiary structural changes can be monitored by CD. However, there are certain limitations in studying the conformational changes using CD especially in the near-UV CD region. These are: (a) conformational changes can be seen only if the internal asymmetry of the aromatic residues is altered due to ligand-binding; (b) it is often times difficult to assign bands to chromophores due to their overlapping; and (c) usually high concentrations of the protein are needed.

In contrast, fluorescence spectroscopy has been found to be one of the most versatile and sensitive methods for probing the tertiary and quarternary structures of proteins. Protein fluorescence in the UV region is mainly due to tryptophan and, to a lesser extent, to tyrosine residues (Konev, 1967; Burstein *et al.*, 1973). Using the differences in the absorption spectra of these two amino acids, one can choose conditions of fluorescence excitation under which protein fluorescence spectra are only due to the tryptophanyl residues (Burstein *et al.*,

1973). Tryptophanyl residues in native proteins occur in different locations and hence are exposed to different microenvironments. The microenvironment of each residue is characterized by a particular set of physicochemical conditions (for example, polarizability, microviscosity, availability of charged groups for interaction with the fluorophore) that influences the fluorescence of the chromophore. Therefore, fluorescence spectroscopy is found to be very useful in studying changes in the microenvironment of a protein due to the interaction with ligands, substrates etc. Moreover, since the parameters of fluorescence are more sensitive to the environment than are those of absorbance and because smaller amounts (10^{-9} to 10^{-12} moles) are easily detected, fluorescence is more frequently used than the difference UV-absorption spectroscopy.

In the present thesis, the emission properties of purified prolylhydroxylase were studied using the Shimadzu spectrofluorphotometer as described in Chapter 2 (section 2.2.4.4). The results are presented below.

The excitation spectrum was first obtained to select the proper wavelength for excitation in subsequent experiments. Figure 4-3 shows the excitation spectrum of native prolylhydroxylase between 250-330 nm recorded at 0 °C. As seen from the figure, the excitation maximum occurs at 285-286 nm. Figure 4-4 shows the emission spectrum of native prolylhydroxylase exciting at 286 nm. The emission maximum occurs at 338-340 nm. The excitation and emission maxima at these wavelengths indicate that the fluorescence properties of prolylhydroxylase are due to the presence of partially exposed tryptophanyl (Trp) residues confirming the suspicion from the CD data, of the presence of Trp. Based on the

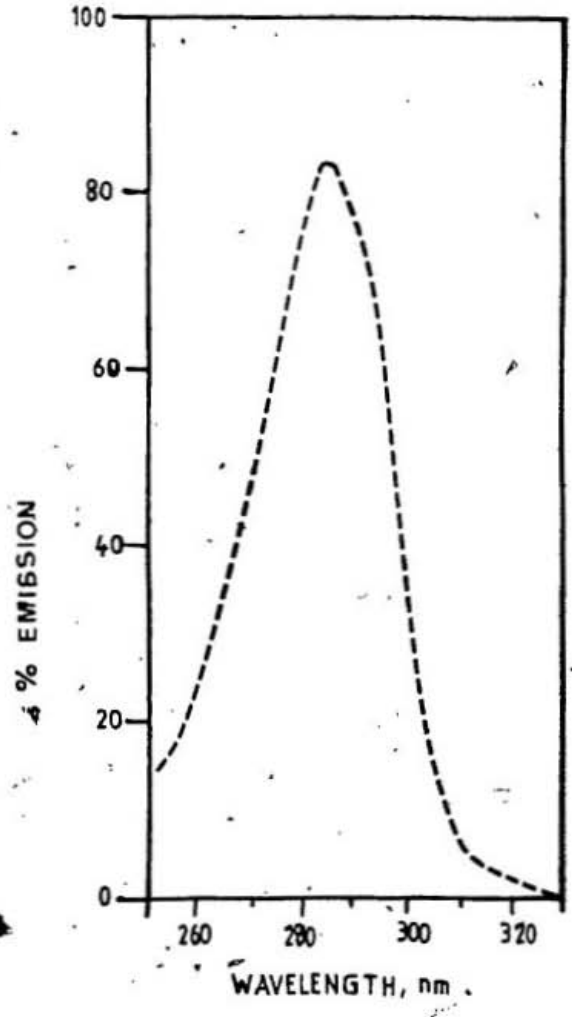


Figure 4-3: Excitation Spectrum of Prolylhydroxylase in Tris-buffer, pH 7.8; concentration 0.02 mg/ml; Temperature $0 \pm 2\,^{\circ}C$; slit width 5nm.

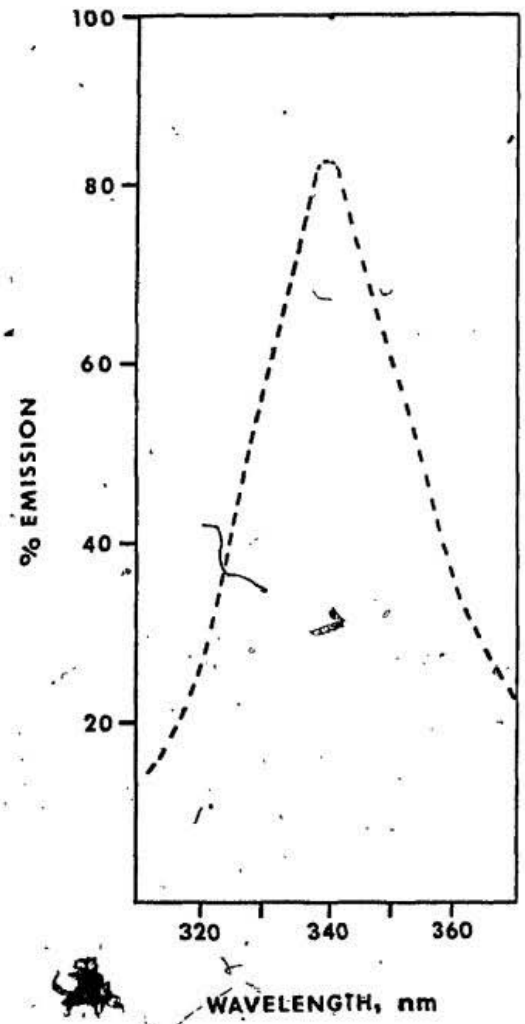


Figure 4-4: Emission Spectrum of Prolylhydroxylase

All conditions as described in the legend for 4-3;

Excitation at 286 nm.

fluorescence spectral properties, Burstein *et al.* (1973) classified Trp residues in proteins into 3 discrete spectral classes. Class I includes Trp residues buried inside the non-polar regions of the protein which usually show spectral maxima in the range of 330-332 nm; Class II includes Trp that are partially exposed to the solvent and exhibit spectral maxima in the range of 340-342 nm; Class III residues are completely exposed to the solvent and are usually on the surface and exhibit spectral maxima in the range of 350-353 nm. Such a classification is however, not a very rigid one, since overlaps are possible. The fluorescence properties of prolylhydroxylase thus can be considered due to Class II (partially exposed) Trp residues.

4.2. Structural Changes in Prolylhydroxylase

Since the function of the enzyme usually depends on its structural integrity as a protein, factors affecting the structure can be useful in monitoring the subtle structure-function relationship. Towards this objective, the effect of temperature on secondary and tertiary structures of prolylhydroxylase was studied using CD and fluorescence techniques. Attempts were made to correlate these changes with its function at the corresponding temperatures. Similar data are not available from the literature.

4.2.1. Effect of Temperature on the Secondary Structure of Prolylhydroxylase

The far-UV CD spectrum was used to monitor the thermal denaturation of prolylhydroxylase. The ellipticity value at 208 nm was selected as an indicator of changes in the α-helix content of the enzyme (see section 6.2.2). Shown in Figure

4-5 is the variation of the fractional α-helical content of the enzyme with temperature. The mid point of the thermal denaturation or melting point (T_m) was found to be 50 °C.

4.2.2. Correlation of Helix-content with the Activity of Prolylhydroxylase

The effect of temperature on the function of prolylhydroxylase was studied at different fixed temperatures and the fractional activity was determined at each temperature. This was then correlated with the structural changes as monitored by changes in the fractional α-helical content with temperature. These results are presented in Figure 4-6. The initial increase in the enzymatic activity between 10-37 °C is apparently due to the temperature effect on the hydroxylation reaction. A fairly steep decline in the activity is however found to occur between 37 °C and 60 °C. As seen from the figure, the fractional α-helical content remained essentially the same between 0-25 °C and slightly decreases at 37 °C indicating a unfolding of the enzyme at this temperature. The observation that the maximal activity occurs at this temperature may imply, in this context, that such slight unfolding of the molecule may be necessary for the enzymatic activity. Both α-helical content and the activity decrease rather sharply after 37 °C indicating the cooperative unfolding of the molecule with the resultant loss of activity. This intimate relationship between the α-helical content and activity of prolylhydroxylase points to the functional importance of these α-helical segments in the enzyme.

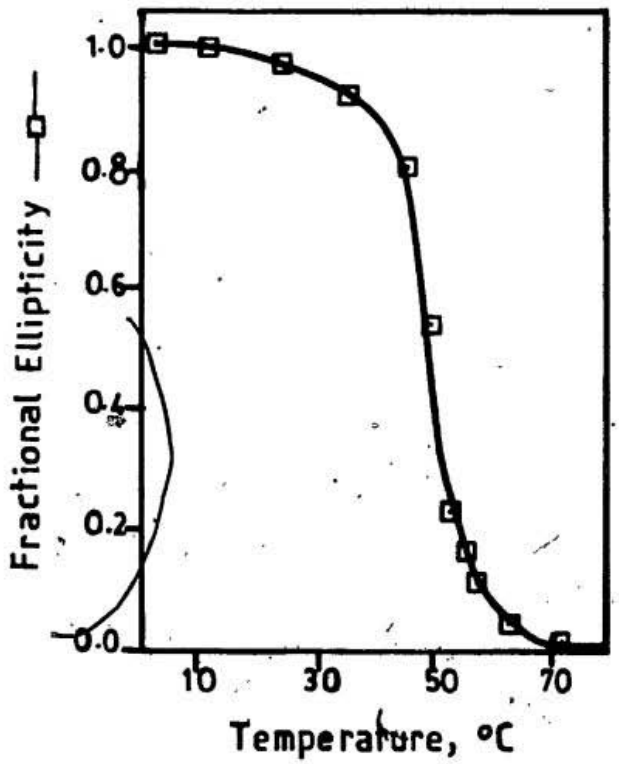


Figure 4-5: Effect of Temperature on the Fractional Ellipticity of Prolylhydroxylase

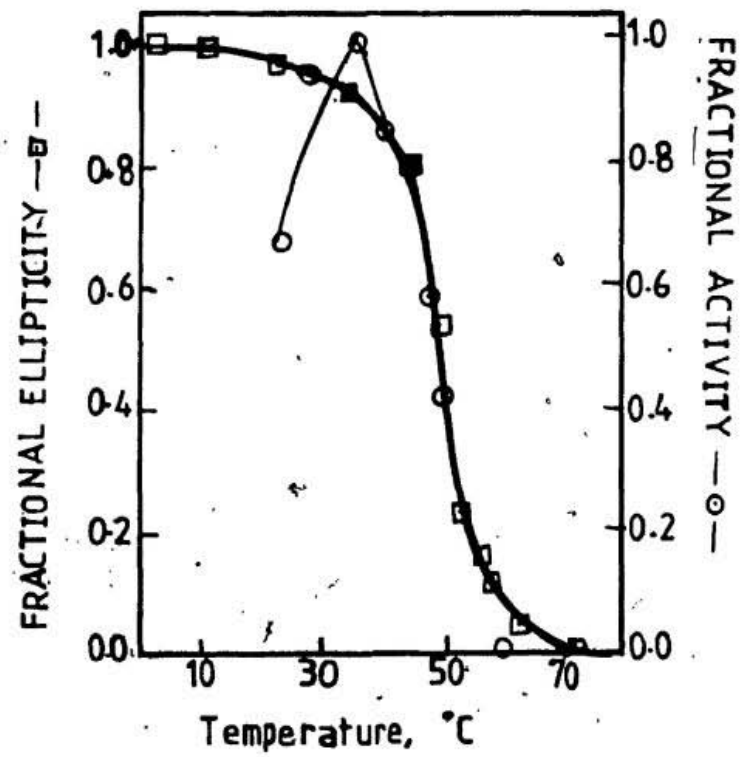


Figure 4-6: Effect of Temperature on the Secondary Structure and Activity of Prolylhydroxylase

Change in the fractional ellipticity (squares) and fractional activity (circles) of prolylhydroxylase as a function of temperature. The fractional ellipticity at 1 °C and fractional activity at 37 °C were taken to be 1.0, while both were taken to be zero at 75 °C.

4.2.3. Effect of Temperature on the Tertiary Structure of Prolylhydroxylase

The perturbations in the asymmetric environment of the aromatic residues are expected to cause changes in the near-UV CD spectra of proteins. The negative CD bands due to the aromatic Cotton effects, in the near-UV CD spectra of prolylhydroxylase vanish completely by heating to over 55-60 °C. This is obviously due to the disruption of the asymmetric environment of these residues in the protein, caused by the collapse of the tertiary structure during the unfolding of the molecule at higher temperatures.

The effect of temperature on the tertiary structure can also be conveniently monitored by fluorescence spectroscopy. Figure 4-7 shows the effect of the temperature on the emission spectrum of prolylhydroxylase. Spectrum "N" represents the native state between 0-4 °C while spectrum "D" represents the spectrum of the denatured enzyme at 90 °C. One can easily see the gradual decrease in the emission intensity due to the increased temperature. Also the emission maximum shifted towards higher wavelengths i.e. from 338-340 nm to 338-347 nm.

The variation of fractional emission as a function of temperature is shown in Figure 4-8. Between 0-10 °C, the emission is unchanged (i.e. 100%) after which one can see a gradual decrease in the emission as the temperature is increased. The mid-point of the transition (50% emission) is at 43.5 °C, about 5 °C lower than that for the decrease in α-helical content. These studies together with CD data indicate that the internal environment and the asymmetry of the enzyme

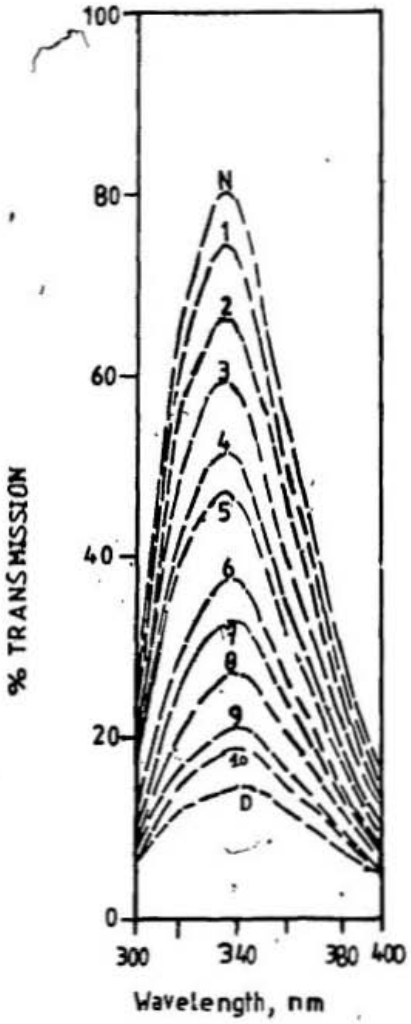


Figure 4-7: Effect of Temperature on the Emission Properties of Prolylhydroxylase

Curve "N" represents the emission spectrum of native enzyme (0 °C) and "D" represents the emission spectrum of the denatured enzyme (90 °C). Spectra 1 - 10 represent those at increasing temperatures between 0 and 90 °C.

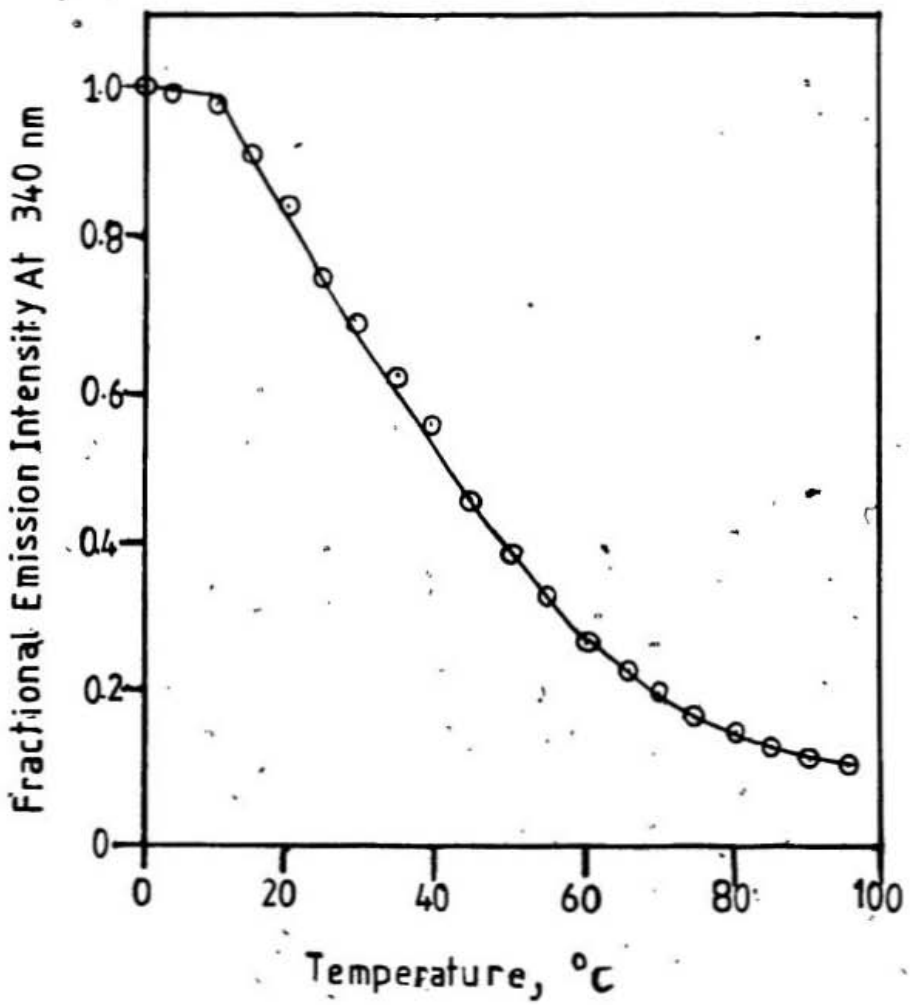


Figure 4-8: Effect of Temperature on the Fractional Emission of Prolylhydroxylase

molecule are altered somewhat earlier than the collapse of the secondary structural elements with increasing temperature.

4.3. Conformational Study of the Interaction of Substrates and Cosubstrates with Prolylhydroxylase

Many enzymes, particularly multi-subunit enzymes, undergo conformational changes upon binding with their substrates and cosubstrates (Citri, 1973). Often, these conformational changes are necessary for the proximal orientation of the reactive groups. It is interesting to study such interactions between prolylhydroxylase and its cosubstrates, cofactors and substrates, in conformational terms which can offer insights into the order of binding of these ligands to the enzyme and the reaction mechanism in general. Conformational changes due to such interactions can be elicited at two levels:

1. at the secondary structural level which may result in gross structural alterations and

2. at the tertiary structural level which may result in rather subtle structural changes.

In the present case, conformational changes due to prolylhydroxylase-ligand interactions were studied using both CD and fluorescence spectroscopy and are described below.

4.3.1. Effects on the Secondary Structure of Prolylhydroxylase

The interaction of prolylhydroxylase with the synthetic substrate, namely, $(Pro\text{-}Pro\text{-}Gly)_5$ was studied in the concentration range of 0.125-1.0 mg/ml of the peptide at 37.5 °C, using far-UV CD spectra (see Methods 2.2.5.2). Similar studies were carried out with the cosubstrate α-KG (0.1-5 mM), with the cofactors $FeSO_4$ (0.05-1.5 mM) and ascorbate (0.5-3 mM). However, no CD spectral changes were detected during the interaction of the enzyme with any one of the above ligands in the indicated concentration ranges. This is taken to indicate that the binding of substrate, cosubstrate or cofactor does not affect the secondary structural elements of prolylhydroxylase.

4.3.2. Effects on the Tertiary Structure of Prolylhydroxylase

The effects of the interaction between prolylhydroxylase substrates or cofactors on the tertiary structure of the enzyme were studied by both near-UV and by fluorescence spectroscopy. No conformational changes were detected by near-UV CD spectroscopy. Since the combined effects of more than one ligand have not been studied at present, it is not known whether the lack of conformational change in the enzyme is due to lack of the correct combination of the ligands. In view of the small amounts of the enzyme available, these studies which require larger amounts of enzyme than the far-UV CD studies, were not pursued further. On the other hand, concentration dependent ligand-binding effects were observed in the fluorescence emission spectra. These interactions are described below.

4.3.2.1. Effect of Substrate-binding on the Emission Spectra of Prolylhydroxylase

The effect of increasing concentrations of the substrate, namely, $(Pro\text{-}Pro\text{-}Gly)_5$ on the emission properties of prolylhydroxylase was studied in the concentration range of 0-2.0 mM by using fluorescence spectroscopy. No changes in the fluorescence spectrum of the enzyme were observed. Similarly, studies using PP-II, a competitive inhibitor of prolylhydroxylase (which was found to have greater affinity towards prolylhydroxylase when compared to the above substrate) also did not elicit any conformational changes in prolylhydroxylase as would be reflected in its emission spectrum. This suggests that substrate binding did not affect the environment of the Trp residues; or it may also be due to the absence of Trp residues near the substrate binding site.

4.3.2.2. Effects of α-KG on the Emission Spectra of Prolylhydroxylase

Figure 4-9 shows the representative spectrum for the effect of α-KG on the emission spectrum of prolylhydroxylase. As seen from this figure, in the presence of α-KG, the fluorescence intensity of prolylhydroxylase is decreased. Figure 4-10 shows the effect of increasing concentrations of α-KG on the relative fluorescence intensity and activity of prolylhydroxylase. The emission data are plotted as the percentage of fluorescence quenching at the emission maximum using the fluorescence intensity of the native enzyme as the reference (100%). As the concentration of α-KG is increased, one can see a gradual increase in the quenching of the emission intensity which saturates after 3 mM α-KG. Shown in the same figure, is the activity of prolylhydroxylase at the same concentration range of α-KG and saturating concentration of $FeSO_4$. The plotted activity is

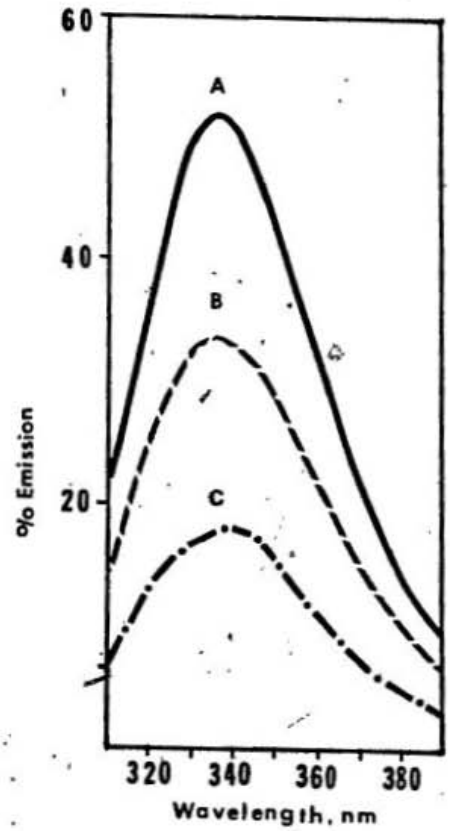


Figure 4-9: Effect of α-KG on the % Emission of Prolylhydroxylase

A = Enzyme alone; B = Enzyme + 5 mM α-KG and

C = Difference spectrum i.e. A minus B

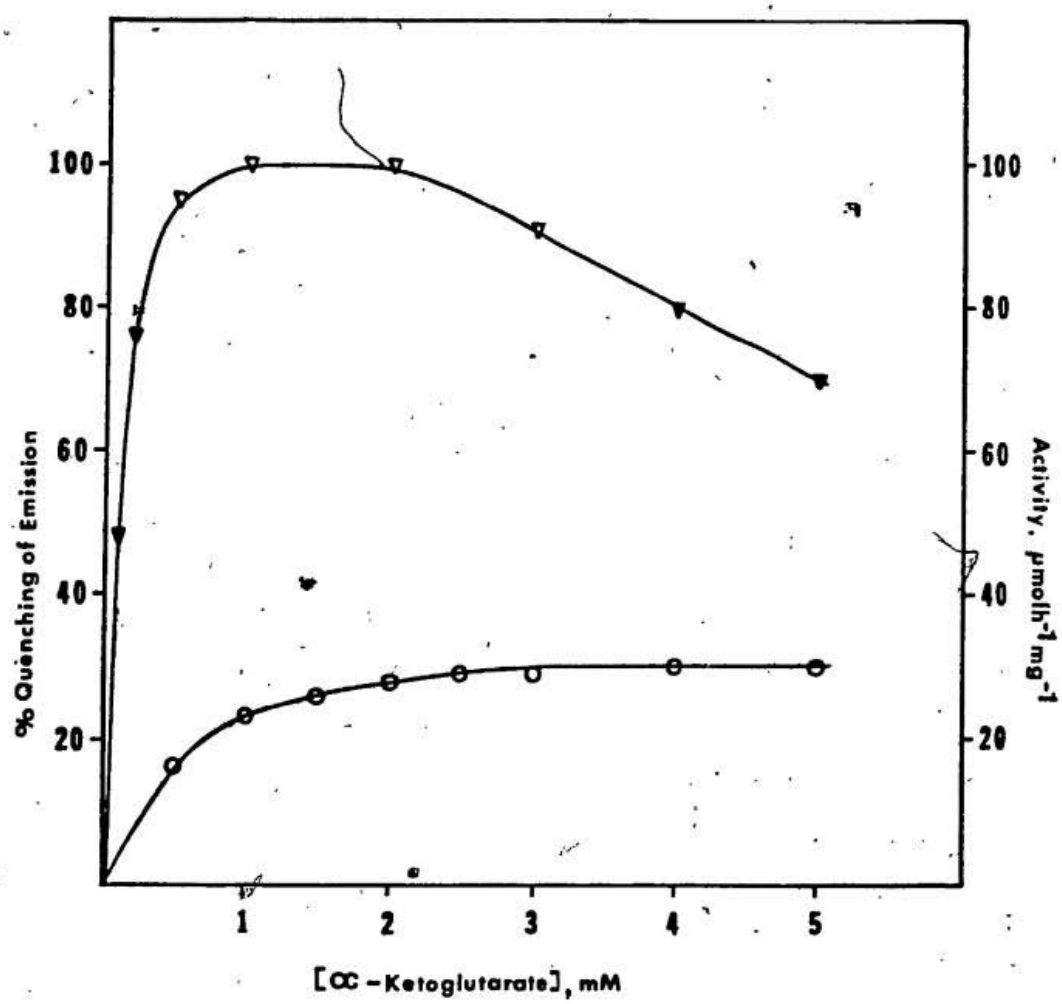


Figure 4-10: Effect of α-KG on the Emission and Activity of Prolylhydroxylase

The % quenching of fluorescence intensity (open circles) was calculated using the fluorescence intensity of the native enzyme as the reference i.e. 100%. The % activity (open triangles) was calculated using the highest activity obtained between 1-2 mM α-KG as the reference (i.e. 100%).

normalized so that the highest activity obtained between 1-2 mM α-KG was taken as 100% and the activity values at other concentrations are adjusted accordingly. As can be seen from the figure, there may be a good correlation between the activity and emission properties of prolylhydroxylase at different concentrations of the cosubstrate. At the maximal activity (at 1-2 mM α-KG), about 20-25% of the emission was quenched; beyond 2 mM α-KG, while the activity starts decreasing, the emission remained more or less constant at 70-75% of the initial value. The alteration in the internal environment as reflected by about 25% quenching may be necessary for the maximal activity under these conditions. The emission maximum of the difference spectrum between that of the enzyme in the presence and absence of α-KG seems to exhibit a slight shift towards higher wavelengths i.e. from 338-340 nm to 338-347 nm.

4.3.2.3. Effect of $FeSO_4$ on the Emission Spectra of Prolylhydroxylase

Figure 4-11 shows the representative spectrum for the effect of $FeSO_4$ on the emission spectrum of prolylhydroxylase. As seen from this Figure, in the presence of $FeSO_4$, the fluorescence intensity of prolylhydroxylase is decreased. Figure 4-12 shows the effect of increasing concentrations of $FeSO_4$ on both the emission and activity of prolylhydroxylase. The emission changes are plotted as the percentage quenching of fluorescence intensity *versus* the concentration of $FeSO_4$. The percentage quenching of fluorescence intensity was calculated using the fluorescence intensity of the native enzyme as the reference (100%). The activity values shown are normalized so that the highest activity obtained between 0.04-0.05 mM concentration of $FeSO_4$ was considered as the reference (100%). The concentration of the cosubstrate was saturating under these

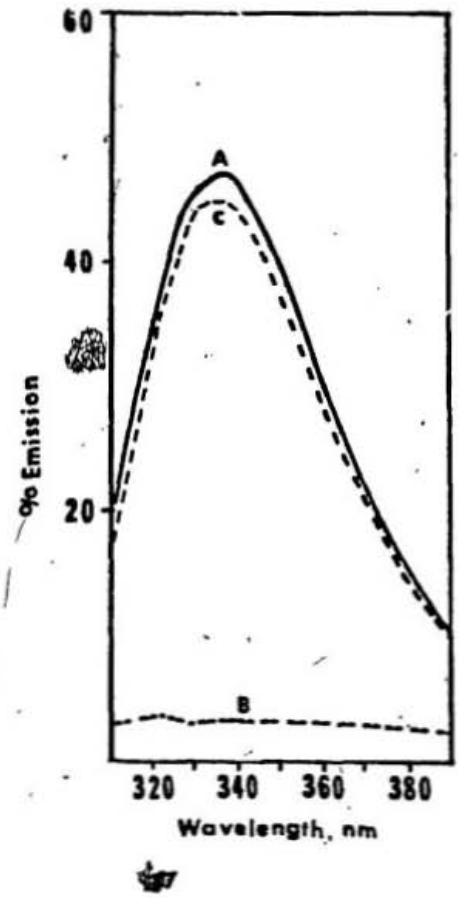


Figure 4-11: Effect of $FeSO_4$ on the % Emission of Prolylhydroxylase

A = Enzyme alone; B = Enzyme + 250 μM $FeSO_4$

C = Difference spectrum i.e. A minus B

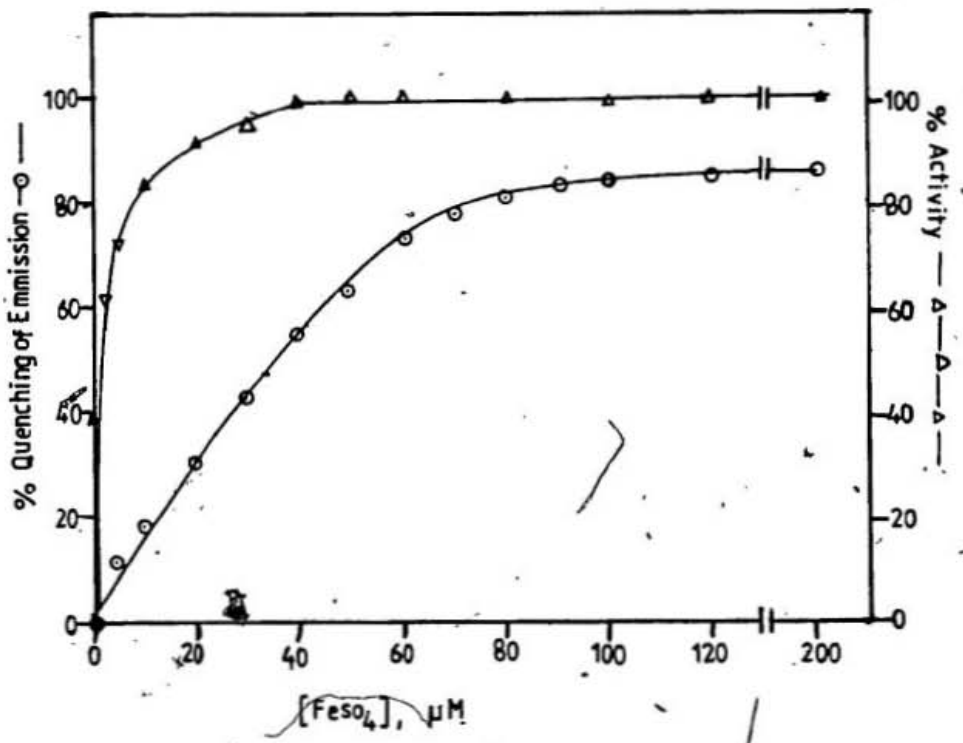


Figure 4-12: Effect of $FeSO_4$ on the Emission and Activity of Prolylhydroxylase

The % quenching of fluorescence intensity (open circles) was calculated using the fluorescence intensity of the native enzyme as the reference (100%); the % activity(open triangles) was calculated using the highest activity obtained between 0.04 - 0.05 mM $FeSO_4$ as the reference (100%).

conditions . As seen from the figure, as the concentration of $FeSO_4$ increases, the intensity at the emission maximum is quenched linearly until 0.07 mM after which one can see a saturation effect. There seems to be a clear correlation between the changes in the fluorescence intensity and activity at different concentrations of $FeSO_4$. At the maximum activity (i.e. at 0.04-0.05 mM $FeSO_4$), the tryptophanyl emission is quenched to about 50-55% of the initial value. This indicates the changes in the internal environment around these Trp residues are due to the binding of $FeSO_4$. However, since no shift is observed in the emission maximum, it suggests that the polarity of the environment remains the same, although the excitation energy of the tryptophanyl residues is reduced due to the binding of the cofactor, namely, $FeSO_4$.

4.4. Discussion

A crucial problem in the field of enzyme chemistry and mechanism concerns the structural factors determining the specificity of enzymes towards their physiological targets. The specificity of an enzyme could originate from the active site geometry of the enzyme itself, it being able to accommodate the substrate of only a particular size, shape or conformation. Such an enzyme is described having a fixed or template-type of active site which demands certain structural dictates in the substrate. Alternatively, the specificity could also be achieved by the flexibility of an enzyme, which can undergo specific conformational changes upon binding of the substrate; these are called substrate-induced conformative responses (Citri, 1973). Such flexibility and movement of a particular domain or segment near the active site seem to be a theme in enhancing the enzyme's specificity and recognition of proper substrates as demonstrated in many enzymes (see Citri, 1973; Jencks, 1975; Koshland, 1976).

The available data on the active site geometry of prolylhydroxylase are rather limited and the conformational changes, if any, upon binding of the substrates, cosubstrates or cofactors have not been studied at all. Therefore, in order to understand the specificity of prolylhydroxylase in terms of the enzyme structure itself, an attempt has been made in the present Chapter, to study the structure and interactions of the enzyme with its cofactors, cosubstrates and substrate, in conformational terms.

The spectroscopic data presented in this Chapter, pertaining to the structural characteristics of purified prolylhydroxylase provide, to my knowledge, the first indications of the conformational features of this enzyme. From the far-UV CD data, prolylhydroxylase seems to be relatively rich in α-helix which comprises as much as 40% of the total secondary structural content. During the writing of this thesis, the β-subunit of human prolylhydroxylase has been sequenced by molecular cloning (Pihlajaniemi *et al.*, 1987) and shown to be the product of the same gene that codes for protein disulphide isomerase (PDI) (E.C 5.3.4.1). The proteins exhibit about 94% sequence homology at the amino acid level and about 84% at the nucleotide level. Edman *et al.* (1985) have studied PDI and shown the presence of substantial α-helical regions in this protein. In view of the similarity between PDI and β-subunit, the latter is also expected to be very α-helical. The tetrameric prolylhydroxylase ($\alpha_2\ \beta_2$) was found to be rich in α-helix and therefore, it may be possible that the β-subunit makes a major contribution to the enzyme's α-helical content.

Although the amino acid composition data (Berg *et al.*, 1979; Kivirikko and

Myllyla, 1982) of prolylhydroxylase available so far does not indicate the presence of Trp residues, the sequence data on β-subunit mentioned above indicates the presence of 5 Trp residues (Pihlajaniemi *et al.*, 1987). In the present study, the near-UV CD spectra of prolylhydroxylase indicated the presence of asymmetrically oriented Trp residues (see section 6.1.1.). This is further confirmed by the fluorescence spectrum which exhibited the excitation maximum at 286 nm and emission maximum at 338-340 nm. As mentioned earlier, these spectral maxima are due to partially exposed (class II) Trp residue (Burstein *et al.*, 1973). In the light of the above observations, it appears likely that some of the secondary structural features and fluorescence properties of prolylhydroxylase are contributed by the β-subunit. Since such data are not available on the α-subunit of prolylhydroxylase, its contribution to the structural features of prolylhydroxylase is not known at present.

The effect of temperature variation on the enzyme's conformation and activity clearly suggests the functional importance of the α-helical secondary structure in the case of prolylhydroxylase.

The absence of any secondary structural changes in the enzyme due to the interaction by substrate, cosubstrate and cofactor is indicative of the lack of movement of major domains or polypeptide segments upon the binding of the above ligands and the relative rigidity of the secondary structural frame work of the enzyme near active site. However, the internal environment around the active site seems to be affected due to the binding of the cofactor FSO_4 and cosubstrate α-KG as revealed by the concentration-dependent changes in the fluorescence

spectra of the enzyme. These changes also correlate well with the activity of prolylhydroxylase at these concentrations of $FeSO_4$ and α-KG. A careful analysis of the above observations provides insights into the active site geometry of prolylhydroxylase and these are summarized below:

1. The changes in the emission spectra indicate the presence of Trp residues near the active site.

2. The active site is already created i.e. no major structural changes are necessary to create an active site to accommodate the incoming substrate molecule. This indicates a rather fixed active site geometry of prolylhydroxylase.

3. The substrate-binding site and the actual catalytic site are separate from each other. For example, while the catalytic site is in the interior, the substrate-binding site may be on the surface thereby substrate-binding does not affect the internal environment.

4. An alternate possibility is that cosubstrate and cofactor binding precede the substrate-binding during the actual reaction. The changes brought about by the cofactor and cosubstrate interactions may be sufficient so that the subsequent binding of the substrate to the enzyme may not result in any further changes.

4.4.1. Analysis of the Substrate-binding Site of Prolylhydroxylase

Many previous studies have demonstrated that the tetrameric structure is necessary for prolylhydroxylase activity, whereas free α and β subunits are inactive (Cardinale and Udenfriend, 1974; Prockop *et al.*, 1976; Kivirikko and Myllyla, 1980). Earlier studies by Kivirikko *et al.* (1971) and Berg *et al.* (1977) have demonstrated that the different tripeptide units in $(Pro\text{-}Pro\text{-}Gly)_5$ and $(Pro\text{-}Pro\text{-}Gly)_{10}$ were hydroxylated to different extents and the penultimate tripeptide from the N-terminus was the most hydroxylated. These observations were explained by a model where prolylhydroxylase has an asymmetric active site in which binding subsites are located adjacent to, but not symmetrical with the catalytic subsite (Berg *et al.*, 1977).

More recently, de Waal *et al.* (1985) have demonstrated the presence of possible substrate-binding subsites in prolylhydroxylase. These authors first covalently blocked the active site by photoinhibition, using a photoaffinity label, N-(4-azido-2-nitrophenyl)-Gly-$(Pro\text{-}Pro\text{-}Gly)_5$ and then checked whether the inactive enzyme still bound to the polyproline affinity column. The covalently-bound photoaffinity label did not impair the binding of the enzyme to the polyproline column although it inhibited the hydroxylation of the synthetic substrates suggesting the presence of binding subsites for substrates. These affinity label studies also indicated that the substrate-binding site and catalytic site are separate from each other, as seen also from the fact that the uncoupled decarboxylation of α-KG catalyzed by the enzyme was not affected by the presence of the affinity label. Moreover, the labelling was specific for the

α-subunit of the tetrameric $\alpha_2\beta_2$ enzyme indicating that the substrate-binding site (s) is (are) present on the α-subunit.

It has been shown tha heat-inactivated prolylhydroxylase still retains the ability to bind to the polyproline affinity column (Ananthanarayanan, unpublished results). This suggests that although the catalytic site is disrupted by thermal denaturation, the substrate-binding site is not affected indicating a stable and defined substrate binding site, probably on the surface of the enzyme, in both the native and denatured states. Thus, binding of the substrate to native prolylhydroxylase would not be expected to result in changes either at the tertiary or at the secondary structural level, if the substrate binds at a pre-created site on the surface of the enzyme, distinct from the catalytic site. This is in agreement with the observations made earlier in this Chapter.

The alternate possibility that the changes brought about by the cofactor and cosubstrate interactions may be sufficient for subsequent binding of the substrate to the enzyme would indicate an ordered or sequential mode of reaction mechanism for proline hydroxylation. The results of extensive kinetic studies (Tuderman *et al.*, 1977 a,b; Myllyla *et al.*, 1977, 1979) and other data (Nietfield *et al.*, 1982; De Jong *et al.*, 1982) are consistent with the binding of Fe^{+2}, α-KG and the peptide substrate to the enzyme taking place in this order. However, the order of release of products is unknown.

4.4.2. Analysis of the Catalytic Site of Prolylhydroxylase

The effect of cosubstrate and cofactor binding on the fluorescence spectra of prolylhydroxylase indicates subtle changes taking place in the internal environment of the enzyme molecule due to this interaction. These changes can be explained if the catalytic site is assumed to be at the interior of the molecule and the binding sites for the cosubstrates and cofactors are either inside or very near the catalytic site and hence their binding affects the internal environment. Available data on cosubstrate binding seem to be consistent with the above assumption.

According to the stereochemical mechanism proposed by Hanauski-Abel and Gunzler (1982), the active site ferrous ion is situated in a pocket of the catalytic subunit. This pocket also accommodates the locus for the decarboxylation of α-KG and subsequent formation of succinate so that the α-KG binding site is proximal to the ferrous ion binding site. They also proposed that the molecular oxygen binding site is situated within this pocket whereas the substrate binding site is on the outside of this pocket. Earlier evidence for the spatial separation of catalytic and substrate-binding sites came from the studies in which specific antibodies to the enzyme strikingly reduced the hydroxylation, but did not affect the decarboxylation (Counts *et al.*, 1978). This was explained to result from the location of the substrate-binding site being readily accessible to the antibodies, while the α-KG binding site is inaccessible since it is situated inside the molecule.

A more direct clue to the nature of α-KG-binding site is obtained from the studies by Majamaa *et al.* (1984). The structure and function of the

α-KG-binding site was studied by these authors by assaying the inhibitory potential of 24 selected aliphatic and aromatic compounds that are structurally analogous to the cosubstrate, α-KG. Based on these data, Majamaa *et al.* (1984) proposed that the α-KG-binding site can be divided into 3 distinct subsites: Subsite I is probably a positively charged side chain of the enzyme that ionically binds the C_5 carboxyl group of α-KG; Subsite II consists of two cis-positioned equatorial coordination sites of the enzyme-bound ferrous ion and is chelated by the C_1-C_2 moiety, while subsite III involves a hydrophobic binding site in the C_3-C_4 region. These authors also point out the importance of subsite I in providing the proper alignment of the cosubstrates which may facilitate subsequent binding of molecular oxygen to the Fe^{+2} at the active site via allosteric rearrangement of the enzyme protein. Later studies by the same authors (Majamaa *et al.*, 1985) also suggested a considerable degree of protein flexibility at the α-KG-binding site, as this site can harbour molecules with markedly different structures and volumes. Shown in the Figure 4-13 is the active site of prolylhydroxylase as originally proposed by Hanauski-Abel and Gunzler (1982) and later slightly modified by Majamaa *et al.* (1984).

Using various hydroxybenzoates and hydroxybenzenes and related compounds that resemble structurally both α-KG and ascorbate, Majamaa *et al.* (1986) have further reported that these compounds inhibit prolylhydroxylase competitively with respect to both cosubstrates. They proposed that the ascorbate binding site is partially identical with that of α-KG-binding subsite II and consists of two cis-positioned coordination sites of the enzyme-bound Fe^{+2} ions. This mode of interaction suggests that ascorbate reduces the enzyme-bound iron through an "inner sphere" mechanism.

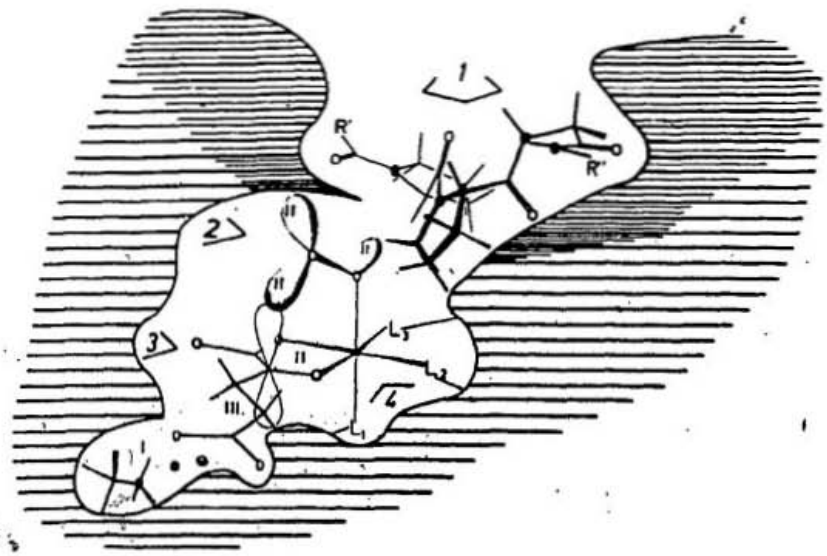


Figure 4-13: Active Site of Prolylhydroxylase

Binding of the peptide substrate, ferrous ion and cosubstrates at the active site; The iron (4) is thought to be located in a pocket and the α-KG (3) to become bound via subsites I, II and III. Molecular O_2 is thought to be bound end-on in an axial position, yielding a superoxide-like structure. The peptide substrate (1) is sterically oriented to participate stereospecifically in the hydroxylation reaction (From Majamaa *et al.*, 1984).

The latest studies by de Waal *et al.* (1987) using again a photoaffinity label, 5-azido-pyrydine-2-carboxylic acid and by Gunzler *et al.* (1987) using coumalic acid (2-oxo-1,2, H-pyran-5-carboxylic acid) confirm that the α-KG-binding sites of prolylhydroxylase are located within the α-subunit.

Based on the above data, the active site geometry of prolylhydroxylase can be visualized as follows : The substrate-binding site and catalytic site are separate from each other. The substrate-binding site is on the exterior (probably on the surface) and the catalytic site is at the interior, inside a pocket which harbours both α-KG binding site and ferrous ion binding site. Molecular oxygen also binds inside this pocket. The substrate-binding site and α-KG binding site are present on the α-subunit of the enzyme while both the subunits contribute to the formation of catalytically active, active site (Gunzler *et al.* unpublished data quoted in Pihlajaniemi *et al.*, 1987).

4.4.3. The β-subunit and its Contribution to Proline Hydroxylation

The contribution of the β-subunit to the catalytic site of prolylhydroxylase was revealed by studies (Hoyhtya *et al.*, 1984) in which a monoclonal antibody to the β-subunit inhibited prolylhydroxylase activity. However, the actual mode of involvement of this subunit in proline hydroxylation is not known at present. The very recent sequence studies by Pihlajaniemi *et al.* (1987) showed that the β-subunit of prolylhydroxylase is in fact, a protein disulphide isomerase (PDI), imparting isomerase activity to prolylhydroxylase also (i.e. prolylhydroxylase exhibits dual enzymatic functions). Structural studies on PDI by Edman *et al.* (1985) showed that the molecule contains two pairs of regions with internal

homology. Two of these regions (a and a') exhibit a highly conserved stretch of 16 amino acid residues with only one amino acid difference between the counterparts. These sequences are found to be marked by the presence of two Cys residues separated by a Gly and His and form the active site of PDI. This agrees with my observation of the pH dependence of prolylhydroxylase activity, suggesting the involvement of His in the active site. These regions of internal homology with two Trp-Cys-Gly-His-Cys-Lys sequences were found to be identical in the human β-subunit. The signal sequences were also highly homologous, although this region in rat PDI was found to be two amino acid residues longer (Pihlajaneimi *et al.*, 1987). Rat PDI and hence β-subunit also share sequence homology with bacterial thioredoxin. According to Pihlajaneimi *et al.* (1987), bacterial thioredoxins (Gleason *et al.*, 1985) were found to be homologous to amino acid residues 6-113 and 349-453 of regions a and a' in the human β-subunit; conservation of amino acid sequences was found along the whole thioredoxin sequence and the corresponding two regions of the β-subunit, but the highest homology was seen in the regions of the two active site Cys residues. These findings suggest a common origin for bacterial thioredoxin, rat PDI and parts of the β-subunit of human prolylhydroxylase. Analogous to thioredoxin, PDI catalyzes the rearrangement of disulphide bonds using the two pairs of Cys residues present in its active site.

Earlier studies with prolylhydroxylase in the presence of p-chloromercuribenzoate (Halme *et al.*, 1970) and N-ethylmaleimide (Popenoe *et al.*, 1969) have suggested that SH groups play a role in proline hydroxylation. Hobza *et al.* (1973) proposed that the ferrous ions are liganded to the enzyme through

the SH side chains of Cys residues and are very close to the α-KG-binding site. Besides their participation in actual hydroxylation, the SH groups are also shown to link the enzyme subunits (Berg *et al.*, 1979). The function of the β-subunit which exists in excess of the α-subunit in the cell has not been defined so far (Kivirikko and Myllyla, 1982). Therefore, the recent observation that the β-subunit itself is an enzyme with disulphide isomerase activity came as an unexpected finding (Pihlajaniemi *et al.*, 1987). Nevertheless, the disulphide rearrangement of prolylhydroxylase was clearly discussed by Nietfield and Kemp (1981). They showed that the enzyme gets inactivated when preincubated in the presence of α-KG and O_2 and this inactivation leaves the enzyme in a tetrameric structure which is different from that of the native structure. This inactivation could be reversed by dithiothreitol (DTT) but not by ascorbate, although both offer protection against this inactivation. It appears that, in the presence of α-KG, O_2 causes not only the oxidation of Fe^{+2} to Fe^{+3} which can be reversed by ascorbate, but also of other groups that can be reversed by DTT but not by ascorbate. Circumstantial evidence indicates that these groups could be SH groups which are rearranged within the enzyme. Presumably, pre-incubation with α-KG induces a conformational change conductive to the oxidation of these SH groups resulting in the mismatched S-S bridge formation. These authors also explained that DTT with a E'_o of --332 mv (Cleland, 1964) can reduce the S-S bonds (E'_o = --200 mv) while ascorbate with E'_o of +60 mv cannot reduce these bonds and reactivate the enzyme. However, ascorbate could protect the enzyme by keeping the SH groups reduced in an indirect way, possibly by preventing the conformational change necessary for inactivation.

In retrospect, these SH group rearrangements might be considered to be due to the isomerase activity of prolylhydroxylase itself, through its β-subunit which could be triggered by the binding of α-KG, Fe^{+2} and molecular oxygen to the enzyme catalytic site in the absence of the substrate. Koivu and Myllyla (1986) and Koivu *et al.* (1987) showed that the tetrameric prolylhydroxylase retains 50% of the isomerase actvity of the isolated β-subunit or PDI itself. According to Koivu and Myllyla (1986), PDI or β-subunit may be an enzyme that is able to retain the tetrameric prolylhyroxylase structure in its native conformation, in which it has maximum enzymatic activity, but it seems to lack the ability to associate the enzyme monomers into the active tetrameric form.

4.4.4. Active Site Geometry in Collagen Hydroxylases

Since the individual binding sites for the peptide substrate, cosubstrates (α-KG and O_2), cofactors (Fe^{+2} and ascorbate) are identified in prolylhydroxylase as discussed above, it would be interesting to characterize the structural features of the closely related enzymes like prolyl-3-hydroxylase, lysylhydroxylase and the distantly related prolylhydroxylase from plants (Chrispeels, 1984). While prolyl-4-hydroxylase binds specifically to both PP II and collagen, prolyl-3-hydroxylase does not bind to PP II but binds collagen. Similarly, lysylhydroxylase binds to collagen but not to PP II (Kivirikko and Myllyla, 1982). On the other hand, plant prolylhydroxylase binds to both PP II and collagen similar to prolyl-4-hydroxylase (Chrispeels, 1984). These observations indicate that there may be differences in the substrate-binding sites or subsites among these enzymes. Kraut (1977) observed a similar situation in the case of serine proteases which catalyze similar reactions and hence show similar spatial arrangements of chemically important

groups but show differences in the binding subsites for different substrates. Thus, it appears that the further refinement of specificity within a family of enzymes can be introduced within the realm of the active site that is common to every member of the family, by introducing binding subsites.

α-KG is utilized by not only prolyl-4-hydroxylase, prolyl-3-hydroxylase and lysylhydroxylase but also by α-KG dehydrogenase (E.C. 1.2.4.2). The last enzyme, however, differs distinctly from the collagen hydroxylases with respect to other cosubstrates and reaction mechanism (Reed, 1974). Studies by Majamaa *et al.* (1985) using both aromatic and aliphatic analogues of α-KG suggest that all the above enzymes including α-KG dehydrogenase have very similar but not identical α-KG binding sites. The subtle differences appear to be again at the subsites which, in turn, define the varied specificity of these enzymes towards their cosubstrate, namely, α-KG.

In the light of the above observations, it appears that the active site geometry of collagen hydroxylases might have been evolved from a common ancestor with slight variations in the theme. However, extensive sequence data as well as X-ray studies on the crystals of these enzymes are needed to provide further insights into the evolution of these enzymes.

4.5. Conclusions

Based on the above discussion, the specificity of prolylhydroxylase appears to be defined at two levels :

1. The specificity of the enzyme towards the substrate is governed by the

structural dictates of the substrate-binding site, which is already defined on the enzyme. In this case, the enzyme demands certain structural features in its substrates for optimal binding and productive interaction.

2. On the other hand, the specificity of interaction of the cosubstrates and cofactors with prolylhydroxylase appears to be governed by the flexibility of the enzyme i.e. the enzyme has to make the necessary changes in its structure for the productive interaction with the ligands. The subtle conformational changes observed upon binding of α-KG and $FeSO_4$ to enzyme would support this argument.

Chapter 5

Spectral Characterization of Peptide Structure

5.1. Methods of Peptide Characterization

Many methods exist that can yield information about the conformation of peptides. These include X-ray crystallography for the solid-state structure and for solution-state conformation, Nuclear magnetic resonance (NMR), IR, CD and fluorescence spectroscopy. In addition, potential energy calculations on peptide conformation provides another useful technique. Among these, the most definitive is X-ray diffraction. However, the drawback of X-ray studies is that the molecules are viewed in conformation (s) determined by the inter- and intra-molecular crystal lattice forces. Still, the observed conformation(s) can always be counted among those adopted by the peptide and, in many instances, will be closely related to conformation(s) in other environments.

The solution conformation of peptides is mainly affected by the solvent-solute interactions and hence results in conformational heterogeneity and dynamics. Small linear peptides are usually flexible molecules and undergo conformational interconversions. The amount of time the peptide spends in each of its conformational states and the rate of interconversions must be taken into

consideration in choosing the methods of conformational analysis and in interpreting the data (Rose *et al.*, 1985).

CD is very sensitive to the conformation, since the arrangement of peptide bond chromophores with respect to each other and to the asymmetric fields in the molecule will vary with different conformations. CD has a very fast time scale (10^{-15} sec) (Cantor and Schimmel, 1980). The CD bands arising from π-π^* and n-π^* transitions of the molecule overlap with each other to a significant extent in the peptide chromophore region (Adler *et al.*, 1973; Woody, 1974; Rose *et al.*, 1985). Because of this, CD spectra of peptides usually show the averaged conformations present, in a given set of experimental conditions. The main information that can be easily obtained from the CD is the presence or absence of the ordered structures, especially in different solvents, which makes this technique very popular.

Vibrational spectroscopy, especially IR, has been used as a complementary technique, along with other techniques in determining peptide conformations. The amide vibrations are very sensitive to H-bonding and therefore, have been of particular interest in the analysis of H-bonded or non-H-bonded (free) peptide conformations (Aubry *et al.*, 1979; Boussard *et al.*, 1979; Kopple *et al.*, 1975 and Rao *et al.*, 1980). Like CD, IR also has a fast time scale (10^{-13} sec) and hence one can see IR bands for almost all species present (Rose *et al.*, 1985).

NMR is perhaps the most useful method of peptide conformational analysis in solution. NMR parameters give information about the chemical environments

of nuclei (chemical shifts), geometric relationships between nuclei (coupling constants and nuclear Overhauser effects), accessibility and H-bonding of amide protons (by exchange kinetics and sensitivity of resonances to temperature and solvents) (Jardetzky, 1980; Rose *et. al.*, 1985). Since NMR gives information which is site-specific, the properties of individual residues can be studied. NMR exhibits a time scale of 10^{-2} to 1 sec, which is slower when compared with the peptide conformational interconversion rates (from nano to milli sec). Hence, conformational interconversions yield averaged parameters.

Solid-state NMR methods have been recently developed and found very useful in obtaining peptide conformation in crystals and solid-like environments for example in lipid bilayers (Pease *et al.*, 1981; Frey *et al.*, 1983). These methods complement solution NMR and other techniques in providing information about the conformation of the peptides.

Conformational energy calculations also offer a means of exploring and characterizing the conformational space available to a peptide. This, supplemented by model building and computer graphics will provide information regarding the minimum energy conformations adopted by the peptides, which can be checked against the spectroscopic data.

The strategy used in this thesis is the combined use of CD, IR and model building for arriving at the conformations of the oligopeptides used in the hydroxylation experiments (Chapter 6). Ideally, greater use of NMR has to be made in analyzing the conformations of the peptides. However, a highfield NMR

facility necessary for the analysis of oligopeptide conformation is lacking on this campus. With limited access to the instrument at Dalhousie University, Halifax and elsewhere, it was only possible to obtain NMR data for the tripeptides like t-Boc or N-acetyl Pro-Gly-X, Pro-DAla-LAla and Hyp-Gly-X, and these were found to be very useful in interpreting the present data on Pro-containing oligopeptides. In spite of this limitation, as will be shown later, the conformations arrived at by CD, IR and model building are quite compatible with those reported in the literature either by X-ray or by conformational energy calculations.

In the following sections, the IR and CD spectral features possible for Pro-containing peptides will be presented and discussed in the light of the literature available. A general discussion of the results is then provided.

5.2. The IR Spectra of Peptides

The IR spectra of peptides and proteins exhibit characteristic bands associated with the backbone amide (-CO-NH-) groups (Miyazawa, 1967). Since the conformation of the peptide backbone is usually influenced by the polarity of the surrounding medium, the medium also can cause changes in the characteristic amide frequencies and intensities. Information can thus be sought regarding the conformation of the peptide molecule in different solvents using the amide bands as markers.

There are 9 distinct amide bands : amide A, amide B and amide I-VII all of which arise out of the different vibrational motions of the bonds involved in the amide groups (Miyazawa, 1967; Rao, 1963). Amide A and B primarily originate

from NH stretching vibrations and appear in the region 3100-3500 cm^{-1}. In the amide A region i.e. 3100-3400 cm^{-1}, bands are usually observed for conformations involving both intra- and inter-molecular H-bonds (Rao, 1963; Stimson *et al.*, 1977; Rao *et al.* 1980). The amide B region 3400-3500 cm^{-1} exhibits bands due to "free" NH group, not involved in H-bonding. The relative intensities of these bands are proportional to the ratios of the respective conformations present. Under high resolution, the free NH stretching band of secondary amides appears as a doublet arising from the cis- and trans-isomers (Rao, 1963). Since the proton-donating powers of amides are nearly equal to those of alcohols, the amides usually are involved in NH...OC type of H-bonding forming pseudo-cyclic peptide structures such as the C_5, C_7 or C_{10} structures. The subscripts used in denoting these C_5, C_7 and C_{10} structures represent the number of atoms involved in a pseudo-ring structure connected by intra-molecular H-bonding. These structures are shown in Figure 5-1.

Several investigators have used IR extensively to study the conformation of small peptides and polypeptides (Deber, 1974; Crippen and Yang, 1974; Kopple *et al.*, 1975; Venkatachalapathi and Balaram, 1979; Ananthanarayanan and Shyamasundar, 1981; Brahmachari *et al.*, 1982; Rao *et al.*, 1980, 1983; Hollosi *et al.*, 1985). These studies indicate that, in the amide region, a band at 3360 cm^{-1} is indicative of the C_{10} structure (which is equivalent to the β-turn conformation) which involves a 4 --> 1 (NH...OC) H-bond (Rao *et al.*, 1983; Hollosi *et al.*, 1985). A band at 3330-3340 cm^{-1} is indicative of a C_7 conformation, although inter-molecular association can contribute to this band (to a smaller extent) as

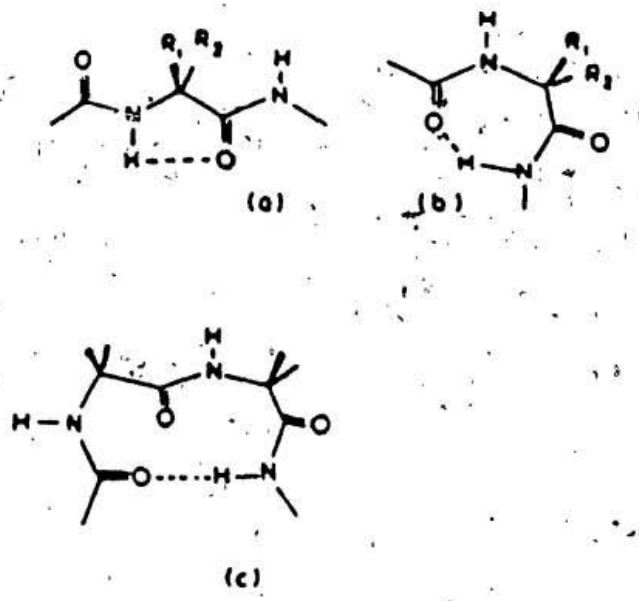


Figure 5-1: Schematic Representation of intra-molecularly H[illegible]ded Structures in Peptides

A = C_5 ; B = C_7 and C = C_{10} strucures

well (Rao *et al.*, 1983). A band in the region 3405 - 3420 cm^{-1} could arise due to a C_5 conformation (Avignon *et al.*, 1969; Maxfield *et al.*, 1979; Rao *et al.*, 1980), although this would be difficult to be distinguished from the free NH vibrations.

The amide group may also be characterized by the other amide bands I - VII. The amide band I is essentially due to carbonyl stretching and appears in the 1670-1700 cm^{-1} region. A band at 1680-1690 cm^{-1} is usually indicative of the peptide carbonyl in a H-bonded conformation (Kawai and Fasman, 1978; Hollosi *et al.*, 1985). A band at 1710 - 1720 cm^{-1} indicates the urethane carbonyl group not involved in H-bonding in the case of t-Boc-protected oligopeptides. Ester carbonyl groups give bands around 1740 cm^{-1}. Tertiary imides like X-Pro bonds show a strong imide band between 1630 - 1670 cm^{-1}, normally appearing at 1650 cm^{-1} (Doyle *et al.*, 1975). Since the tertiary amide bond cannot act as a proton-donor, the presence of this band is also indicative of an open, non-H-bonded peptide backbone. This was observed with PP-II and other Pro-containing polypeptides (Lazarev *et al.*, 1985). The amide bands II and III appear at 1500 - 1550 cm^{-1} and 1200 - 1300 cm^{-1}, respectively, and arise due to the N-H or C-N bending motions. The amide bands IV - VII occur due to the bending and rotational motions of the atoms and appear in the region 700 - 200 cm^{-1}. The latter bands are, however, found to be experimentally difficult to access and to interpret, due to excessive solvent absorption and related problems in these regions (Hollosi *et al.*, 1985). In general, since the amide A, B and I bands usually give the information about the H-bonded and non-H-bonded structures in peptides, they have been extensively used in the characterization of peptide

conformation. In addition, the 1700 - 1760 cm^{-1} region due to ester and urethane carbonyl stretching is also useful in determining small oligopeptide conformations.

Using normal vibrational analysis, Bandekar and Krimm (1986) calculated the amide I, II and III band frequencies for the regular secondary structures like the α-helix, β-sheet and also for different types of β-turns. The considerable agreement between the calculated and experimentally observed frequencies makes them useful in identifying the various types of reverse turns.

5.3. CD Spectra of Peptides

In peptides and proteins, the CD spectra in the far-UV region arise from the peptide bond that absorbs light and undergoes electronic transitions. Two such transitions of importance are, the n - π^* transition and the π - π^* transition (Woody, 1974). The former transition usually results in a CD absorption band in the 220-230 nm region, while the latter gives rise to CD bands farther into the lower wavelength regions extending to the vacuum-UV CD region (up to 150 nm) (Brahms *et al.*, 1977; Brahmachari *et al.*, 1979). The interaction of different transitions may result in the splitting of the absorption band, called the exciton splitting (Bayley, 1980). This is usually observed in periodically ordered structures like the α-helix.

As mentioned earlier, CD has been best used for monitoring the changes in the conformations brought about by different experimental conditions (Brahmachari *et al.*, 1932; Ananthanarayanan and Shyamasundar, 1981; Smith and Pease, 1980). In this section a brief review of the conformations adopted by

the Pro-containing peptides will be presented followed by the description of the CD spectra exhibited by these structures. These are expected to serve as guidelines in interpreting the CD spectra of the test peptides used in the present study.

5.3.1. Conformational Features of Pro-containing Peptides

It has been found that the chiroptical properties of Pro-containing peptides are heavily influenced by prolyl residue contributions (Hollosi *et al.*, 1985). The presence of prolyl residues in a peptide chain is of considerable interest and importance because of the following two characteristics with which they are associated :

1. the geometrically restrictive mobility due to the presence of the pyrrolidine ring as the side chain (with ϕ fixed at about -60°) and

2. the tertiary amino (imino) group cannot participate in H-bonds that normally stabilize the various secondary structures in globular proteins.

Therefore, Pro residues cannot be accommodated in inner α-helix and β-sheet structures which are rather compact and mainly stabilized by recurring H-bonds. The α-helix forming potential for Pro has been found to be very low and, in fact, it has been labelled as a helix-breaker (Chou and Fasman, 1977). However, there are two specific conformations in which Pro residues have been found to be accommodated without any difficulty :

1. an extended, left-handed helix which is stabilized by the stereochemical

restrictions offered by Pro residues alone without being dependent on H-bond formation and

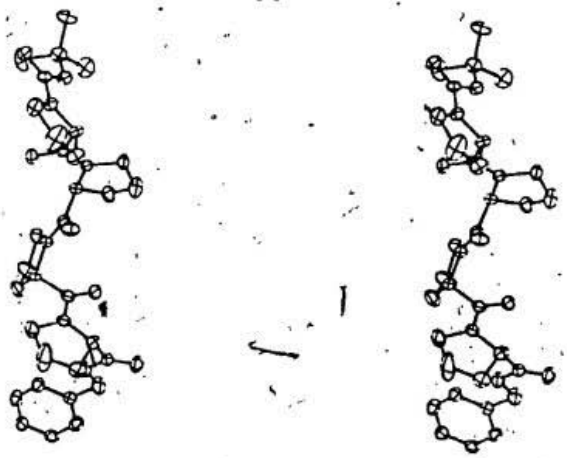
2. β-turn, the folded tetrapeptide sequence which is involved in changing the direction of the peptide chain by 180°, as found in globular proteins.

The above two structures are presented in Figure 5-2.

The first of these structures has been found to be characteristic of the homopolymer of L-proline, poly(Pro) in the form II, with all trans-configuration of the peptide bonds. Similar structures have been observed in the individual chains of collagen and its related homo- and heteropolypeptides (Sasisekharan, 1959a; Brown *et al.*, 1972; Ananthanarayanan *et al.*, 1976). It may be noted that the extended all-trans structure of PP-II can undergo isomerization to yield the all-cis PP-I conformation.

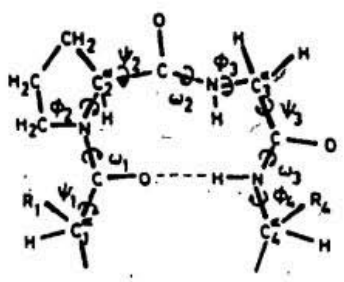
β-turns have been found to be an important class of secondary structures in globular proteins (Smith and Pease, 1980; Rose *et al.*, 1985). A survey of the amino acid sequences found in the turn regions of many globular proteins whose structures are known from X-ray studies clearly indicated that proline is the most frequently occuring bend residue and it prefers the second position of the tetrapeptide sequence (Chou and Fasman, 1977). The third position of the tetrapeptide sequence is found to be preferred by residues like Gly, Asn, Asp, Ser, Thr, Glu and Tyr (Crawford *et al.*, 1973; Chou and Fasman, 1977). A specific analysis of the frequencies of occurence of amino acid residues in Pro-containing tetrapeptide segments in proteins has recently been carried out by Ananthanarayanan *et al.* (1984).

A



Stereo drawing of Boc-Pro_4-BZ molecule
(From Matsuzaki, 1974)

B



A β-turn containing tetrapeptide sequence with Pro in the 2nd and Gly in the 3rd position (From Brahmachari et al., 1981)

Figure 5-2: Conformations of Proline-containing Peptides

A. Left-handed helix of polyproline-II and

B. Type II β-turn with Pro in the 2nd position.

Both the PP-II and β-turn conformations have been found to give distinctive CD spectra and a vast literature is available on the subject.

In view of the above considerations, I have tried to analyze the CD data of Pro-containing peptide substrates of prolylhydroxylase in terms of these two possible conformations, namely, PP-II and β-turns which are likely to exist in equilibrium with the unordered or random-coil structure.

5.4. CD of the Polyproline-II Helical Conformation

The extended left-handed PP-II helical structure of Pro-containing peptides is characterized by a CD spectrum with a small positive band near 225 nm and a large negative band around 200 nm (Tiffany and Krimm, 1969a,b). Most of the studies characterizing this structure were carried out on homo- or heteropolymers, although in a very few cases, oligomers were used. The characteristic CD spectrum of the left-handed extended helix was also observed in the charged forms of poly(Glu) and poly(Lys) although there is no direct evidence in these cases (Tiffany and Krimm, 1969 a,b). A similar CD spectrum was observed for collagen and related polypeptides as shown in the Table 5-1. In many of these peptides, the existence of this conformation was demonstrated in the solid-state by X-ray studies.

From the above studies it appears that, in general, the extended PP-II like left-handed helical conformation is seen when Gly is followed by the Pro moiety in heteropolymers containing these residues. In other studies using oligoprolines, similar CD spectra due to the extended left-handed helical conformation were

Table 5-1: CD Spectral Characteristics of Collagen and Related Polytripeptides

Protein/ Peptide[a]	Conformation in solid-state	CD Maximum nm	CD Minimum nm	Ref.
1. C1q	TH	230	200	1
2. Collagen	TH	220	198	2
3. $(P\text{-}P\text{-}P)_n$	PP-II	226	206	3
4. $(G\text{-}A\text{-}P)_n$	PP-II	-	198	4
5. $(G\text{-}P\text{-}A)_n$	TH	222	198	5
6. $(G\text{-}S\text{-}P)_n$	PP-II	-	198	5
7. $(G\text{-}P\text{-}S)_n$	TH	222	198	5
9. $(G\text{-}L\text{-}P)_n$	TH	-	198	6
10. $(G\text{-}P\text{-}L)_n$	TH	220	201	6
11. $(BA\text{-}P\text{-}P)_n$	TH	230	201	7
12. $(A\text{-}P\text{-}P)_n$	PP-II	225	203	7
13. $(G\text{-}S^*\text{-}P)_n$	Unknown	-	202	8
14. $(G\text{-}P\text{-}S^*)_n$	Unknown	220	198	8
15. $(G\text{-}F\text{-}P)_n$	TH	220	190	9

a amino acids are indicated in one-letter code

TH Triple-helix; PP-II poly(Pro) II; S* Sarcosine; BA β-Alanine.

1. Brodsky-Doyle *et al.*, 1976

2. Brown *et al.*, 1969

3. Tiffany and Krimm, 1969a

4. Doyle *et al.*, 1971

5. Brown *et al.*, 1972

6. Scatturin *et al.*, 1975

7. Bhatnagar and Rapaka, 1976

8. Ananthanarayanan *et al.*, 1976

9. Brahmachari *et al.*, 1978.

observed (Helbecque and Loucheux-Lefebvre, 1978). From the theoretical and experimental studies of Madison and Kopple (1980) and Hollosi *et al.* (1985), it was observed that PP-II like extended conformers also predominate in small Pro-containing linear peptides in polar solvents like water or methanol, due to the dominant chiral contribution of Pro residues (negative band around 200 nm). In addition, the disordered (random-coil) structure of the peptide will also be expected to be present in these solvents. By contrast, folded C_7 or C_{10} confo.mers stabilized by H-bonds are known to prevail in these peptides in order-promoting organic solvents like trifluoroethanol (TFE) or dioxane (Madison and Kopple, 1980; Hollosi *et al.*, 1985). Since the ordered structures will be in equilibrium with the unordered or open conformers in any solution of a peptide in a given solvent, the observed CD spectrum arises from an average of these structures. When the solvent is changed, there will be a redistribution of conformers which will be expressed by a different kind of spectrum.

5.5. CD Spectra of β-Turns

Venkatachalam (1968) carried out a detailed conformational analysis of the tripeptide sequences that can form β-turn structures and classified them into 3 types of β-turn structures I, II and III. Later studies by Lewis *et al.* (1973) further defined additional types of β-turns. However, Woody (1974) was the first to calculate the theoretical CD curves for the various types of β-turns by computing the $n - \pi^*$ and $\pi - \pi^*$ rotational strengths for tripeptides in these conformations. Four classes of CD spectra were proposed for the different types of β-turns (Woody, 1974). Table 5-2 shows these classes of spectra along with the criteria for the classification.

Table 5-2: Classification Scheme for Calculated CD Spectra

Class	Extrema	
	Sign	Wavelength Range, nm
A	-	210-220
	+	195-200
	-	<190
B	-	>220
	+	200-210
	-	<190
C	-	200-210
	+	180-195
D	-	>225
	+	210-220
	-	190-200
	+	<190

After Woody, 1974.

According to Woody (1974), the most common CD spectrum by far is the class B spectrum. Almost all variants of type I, about 80% of type II and about 65% of type III β-turns give CD spectra of the class B pattern. Class B spectra are characterized by a minimum in the region of 220-230 nm and a maximum in 200-210 nm region followed by another minimum below 190 nm. A minor proportion of type II and type III β-turns (about 20%) gives rise to class C spectra which are similar to those of the right-handed α-helix with two negative bands around 220 nm and 205 nm and a positive band around 190 nm. The mirror image of this spectrum, class C' is expected to give positive CD bands around 220 and 205 nm and a negative band below 190 nm resulting in an inverted α-helix pattern. The mirror image of type II β-turn (type II' β-turn) can be associated with class C spectra while class C' spectra can be exhibited by type II β-turns and others like type 13 β-turn (in Venkatachalam's notation) (Smith and Pease, 1980).

Class A spectra are exhibited by about 15% of type III β-turns which are very similar to class B spectra except that the extrema are blue-shifted by about 5-10 nm in class A. Class D spectra are characterized by a positive band between 190 - 200 nm and another positive band below 190 nm. This kind of spectrum is most likely to occur in open β-turns (i.e. β-turn without the 4 --> 1 H-bond).

Although this classification scheme can offer guidelines for characterizing different types of β-turns, the predominance of class B spectra makes it difficult to assign a particular type of β-turn in a given peptide based on CD alone, without the supporting evidence from studies using other physical techniques like IR, NMR and/or X-ray crystallography. Therefore, later investigators utilized the

above-mentioned techniques in addition to CD, in characterizing the different types of β-turns, thus making available reference spectra for various β-turns.

Kawai and Fasman, (1978) have observed Class B spectra in CBZ-Gly-Ser-(OBu^t)-Ser-Gly-O-stearyl ester with a type I β-turn and Crisma *et al.* (1984) in Piv-Pro-Val-$NHCH_3$, with a type I (III) β-turn. Type I β-turn exhibiting a class C spectrum has been reported in cyclo(Gly-Pro-Ala)$_2$. Somewhat similar spectra have been obtained for type I β-turns in Z-Aib-Ala-Aib-Pro-OMe (Crisma *et al.*, 1984) and cyclo(Ala-Ala-ϵ-aminocaproyl) (Bandekar *et al.*, 1982).

In addition to the Pro-containing cyclic and polypeptides, many linear oligopeptides have been shown to exhibit class B spectra due to type II β-turns. Using vacuum-CD, which enables one to record spectra into much lower wavelength regions (165 nm), Brahmachari *et al.* (1979) reported a class B spectrum in N-acetyl-Pro-Gly-Leu-OH and it was attributed to the presence of type II β-turn. In a homologous series of tripeptides with the sequence N-acetyl-Pro-Gly-X-OH (where X = Gly, Ala, Val, Leu, Ile and Phe), Ananthanarayanan and coworkers (Brahmachari *et al.*, 1982) reported type II β-turns which displayed class B spectra. Among these, N-acetyl-Pro-Gly-Phe-OH was demonstrated to adopt type II β-turn in the solid-state by X-ray diffraction (Brahmachari *et al.*, 1981).

Class C spectra have been predicted by Woody (1974) for peptides containing the L-D amino acid sequences. However, it was observed that both

5.6. Conformation of Pro-containing Peptides Deduced from CD and IR Data

The peptides used for IR studies are listed in Table 5-3 along with their parameters, while CD parameters are listed in Tables 5-4 and 5-5. As mentioned earlier in Chapter 2, the concentration of the peptides used for IR studies was 1-5 mg/ml (2×10^{-3} to 10×10^{-3} M) and where possible, 0.5 mg/ml (1×10^{-3}M) was also used. There seems to be practically no inter-molecular aggregation in this concentration range. Rao *et al.* (1980) used a peptide concentration of 5×10^{-3} M in their IR studies and reported no inter-molecular aggregation at this concentration. For CD studies, the peptide solutions were prepared at the concentration of 1-2 mg/ml ($2\text{-}4 \times 10^{-3}$ M). A wider concentration range could not be used due to technical difficulties. The CD spectra of the peptides were obtained in TFE, water and 1,4-Dioxane (dioxane). Spectra in dioxane were taken as an additional aid in interpreting the structure of the two peptides, by comparing their CD with those reported by others in this solvent (Cann *et al.*, 1973; Hollosi *et al.*, 1985). The conformational analysis of the peptides in the present thesis is based on both the CD as well as IR spectra obtained here, and in the light of the information available in the literature on peptides analogous to the present ones using these and other methods.

Table 5-3: ν_{NH} and ν_{CO} for Pro-containing Peptides

Peptide	*ν_{NH} cm^{-1}	+ν_{CO} cm^{-1}
1.Boc-PPG-$NHCH_3$	3450,3360	1740(s),1675,1650(s)
2.Boc-PPGP-OH	3400,3300	1720(s),1680,1600(s)
3.Boc-PPGP-$NHCH_3$	3400,3300	1720(s),1680, 1650(s),1600(s)
4.Boc-PPGPP-OH	3400,3300	1720(s),1680, 1650,1600(s)
5.Boc-PPAP-OH	3410,3310	1720(s),1680(s), 1650,1600(s)
6.Boc-PPQP-OMe	3480,3400, 3320,3180	1740(s),1680, 1645(s),1600(s)
7.Boc-VPGV-OH	3440,3340	1705(s),1685, 1630(s),1600(s)
8.Boc-GVPGV-OH	3420(s),3340	1720(s),1670, 1630(s),1600(s)

* Amide A band

\+ Amide I band

s Shoulder

Table 5-4: Summary of CD Parameters for Pro-containing Tripeptides

Peptide	λ_1, nm	$[\theta]$*	λ_2, nm	$[\theta]$*
1. Boc-PGV-OH				
TFE	225	-3,760	200	+11,750
Water	225	-3,600	200	-10,600
2. Boc-PGL-OH				
TFE	230	-1,670	200	+6,160
Water	230	-1,530	190	-16,000
3. Boc-PPG-$NHCH_3$				
TFE	200	-1,460	198	+510
Water	225	-320	190	-2,545
Dioxane	230	-640	202	+3,180

* $[\theta]$ expressed in $deg.cm^2dmol^{-1}$.

Table 5-5: Summary of CD Parameters for Pro-containing Tetra- and Pentapeptides.

Peptide	λ_1, nm	$[\theta]$*	λ_2, nm	$[\theta]$*
1. Boc-PPGP-OH				
TFE	230	-820	196	-7,200
Water	228	+700	196	-6,625
Dioxane	226	-2,800	204	-5,240
2. Boc-PPGP-$NHCH_3$				
TFE	--	--	198	-11,980
Water	236	+600	198	-8,590
Dioxane	226	-10,180	208	-6,890
3. Boc-PPGPP-OH				
TFE	230	-340	196	-14,410
Water	224	+3,100	196	-17,740
Dioxane	232	-1,690	202	-8,445
4. Boc-PPAP-OH				
TFE	230	-1,320	200	-15,600
Water	226	+1,680	200	-18,120
Dioxane	230	-2,640	208	-6,360
5. Boc-PPQP-OMe				
TFE	232	-480	202	-13,230
Water	226	+1,650	202	-12,950
Dioxane	236	-830	206	-8,270
6. Boc-VPGV-OH				
TFE	228	-1,175	200	+4,400
Water	224	-1,060	190	-4,990
7. Boc-GVPGV-OH				
TFE	224	-1,000	200	+2,100
Water	218	-1,050	194	-3,270

* $[\theta]$ expressed in $deg.cm^2dmol^{-1}$.

5.6.1. t-Boc-Pro-Gly-X-OH

Detailed IR and CD spectral analyses were carried out on the tetra- and higher oligopeptides. Similar conformational information on tripeptides is available to me from the studies reported by Brahmachari *et al.* (1982) on N-acetyl-Pro-Gly-X-OH and the studies on t-Boc-Pro-Gly-X-OH (to be reported by Attah-Poku and Ananthanarayanan). The relevant CD data from the latter study, which were confirmed by the author as well, are presented briefly in this section before the data on the other peptides are presented.

As shown in Figure 5-3, the CD spectrum of t-Boc-Pro-Gly-Leu-OH in TFE at 0 °C exhibits a typical class B spectrum (Woody, 1974). The features of the spectrum include one shallow negative band around 230 nm and a strong positive band around 200 nm. The mean residue ellipticity values at these bands are -1,670 and +6,160 deg.cm^{2}dmol^{-1}, respectively. The CD spectrum in water, on the other hand, displays two negative bands, a shallow one around 230 nm (mean residue ellipticity, -1,530 deg.cm^{2}dmol^{-1}) and a deeper band around 190 nm (mean residue ellipticity, -16,000 deg.cm^{2}dmol^{-1}). The solvent dependence of the CD spectrum indicates the shifts in the conformational equilibria in different solvents.

The CD spectra of t-Boc-Pro-Gly-Val-OH in TFE and in water at 25 °C are also presented in Figure 5-3. In TFE, the spectrum displays a negative band around 230 nm (mean residue ellipticity, -3,760 deg.cm^2.dmol^{-1}) and a strong positive band at 200 nm (mean residue ellipticity, +11,750 deg.cm^{2}dmol^{-1}). In aqueous medium, the spectrum exhibits two negative bands, one around 225 nm (mean residue ellipticity, -3,600 deg.cm^{2}dmol^{-1}) and a negative band around 200

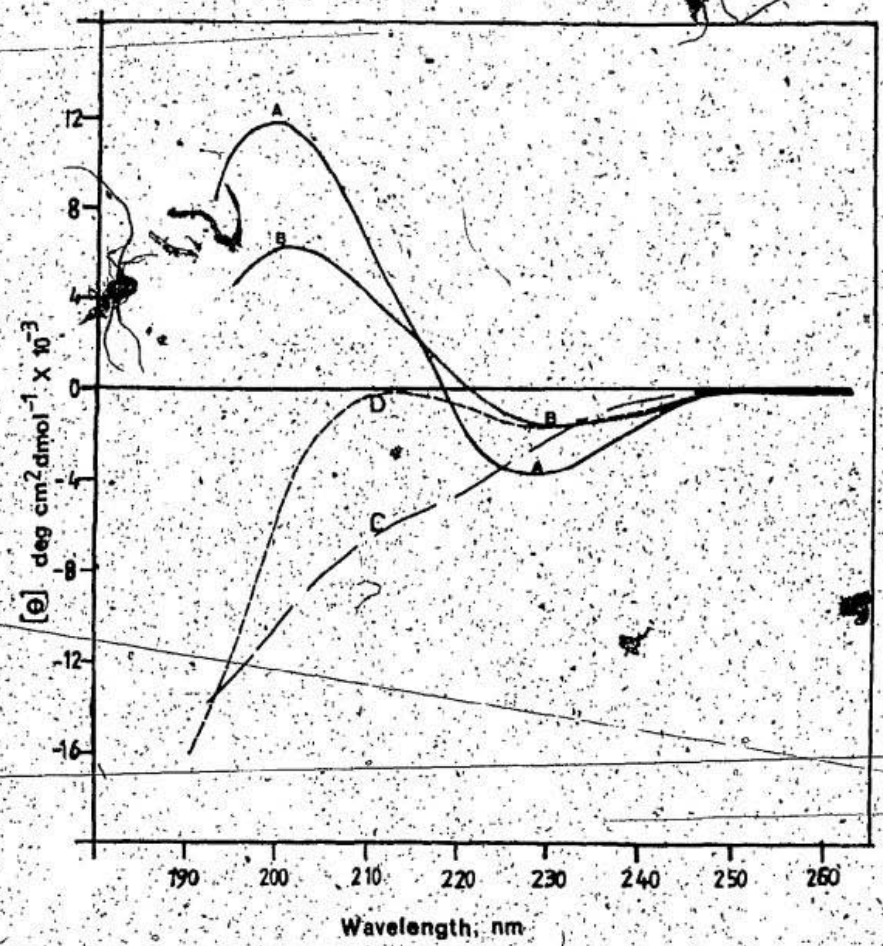


Figure 5-3: CD Spectra of t-Boc-Pro-Gly-Val-OH (1) and t-Boc-Pro-Gly-Leu-OH (2)

A = (1) in TFE; B = (2) in TFE; C = (1) in water and D = (2) in water

Concentration = 1 mg/ml

nm (mean residue ellipticity, -10,600 deg.cm^{2}dmol^{-1}). These features are similar to those of t-Boc-Pro-Gly-Leu-OH as described above.

The presence of class B spectra in TFE indicates the presence of either type I or type II β-turns. Earlier studies on N-Acetyl-Pro-Gly-Leu-OH by vacuum-UV CD demonstrated the presence of a type II β-turn in this peptide (Brahmachari *et al.*, 1979). In fact, the TFE spectrum of N-Acetyl-Pro-Gly-Leu-OH at -40 °C and low peptide concentration (0.2 mg/ml) was reported to be the representative of an isolated β-turn conformation. Further studies by Brahmachari *et al.* (1982) on the homologous series N-Acetyl-Pro-Gly-X-OH (where X = Gly, Ala, Leu, Ile and Phe) using CD, IR and NMR have demonstrated that all the peptides adopt the β-turn conformation, although the stabilities of the β-turn differ in each individual case.

Based on the above studies, the CD spectral features of t-Boc-Pro-Gly-Leu-OH and t-Boc-Pro-Gly-Val-OH can be ascribed to the presence of a type II β-turn conformation. More detailed CD and IR studies were carried out on tetra- and higher oligopeptides that were prolylhydroxylase substrates (Chapter 5). These are presented below.

5.6.2. t-Boc-Pro-Pro-Gly-NHCH$_3$

Earlier studies on Pro-Pro-Gly sequences have shown this to be the minimum peptide sequence hydroxylated by prolylhydroxylase (Prockop *et al.*, 1976). In the light of this observation, it is of interest to study the conformational features of this sequence and hence chosen for the present study using IR and CD.

Figure 5-4 shows the IR spectra in $CHCl_3$, of t-Boc-Pro-Pro-Gly-$NHCH_3$ (Boc-PPG-$NHCH_3$) in the NH (left panel) and CO (right panel) stretching regions. In the NH stretching region, two major bands are present, one at 3450 cm^{-1} and another at 3360 cm^{-1}. The 3450 cm^{-1} band can be attributed to the non-H-bonded (free) NH stretching vibrations and the 3360 cm^{-1} band is due to the H-bonded NH groups. Thus, both H-bonded and free conformers are present. The relative intensities of the bands indicate the predominance of the H-bonded conformers. Boussard *et al.* (1979) and Rao *et al.* (1980) observed that the 3360 cm^{-1} band is characteristic of the 4 --> 1 intra-molecularly H-bonded folded conformation. Therefore, the 3360 cm^{-1} band in the tripeptide indicates the presence of such a folded β-turn conformation. In the carbonyl region, a broad unresolved band appears at 1675 cm^{-1} along with two shoulders, one at 1740 cm^{-1} another at 1600 cm^{-1}. The strong unresolved band tapering at 1675 cm^{-1} also suggests the involvement of the peptide amide in H-bonding. The H-bonded conformers indicated by the IR data of the tripeptide in NH and CO stretching regions, therefore, can be visualized as a C_{10} structure. This C_{10} conformer would form a H-bond between Pro^1 CO and HN of glycyl methylamide (4 --> 1 type).

The CD spectra of Boc-PPG-$NHCH_3$ in TFE, water and dioxane are presented in Figure 5-5. In TFE, a well-pronounced negative band at 220 nm and a positive band at 198 nm are seen. In water, however, there are two negative bands, a shallow trough around 225 nm and another stronger band at 190 nm. In dioxane, a broad negative band is observed around 225 nm and a strong positive band is seen around 200 nm.

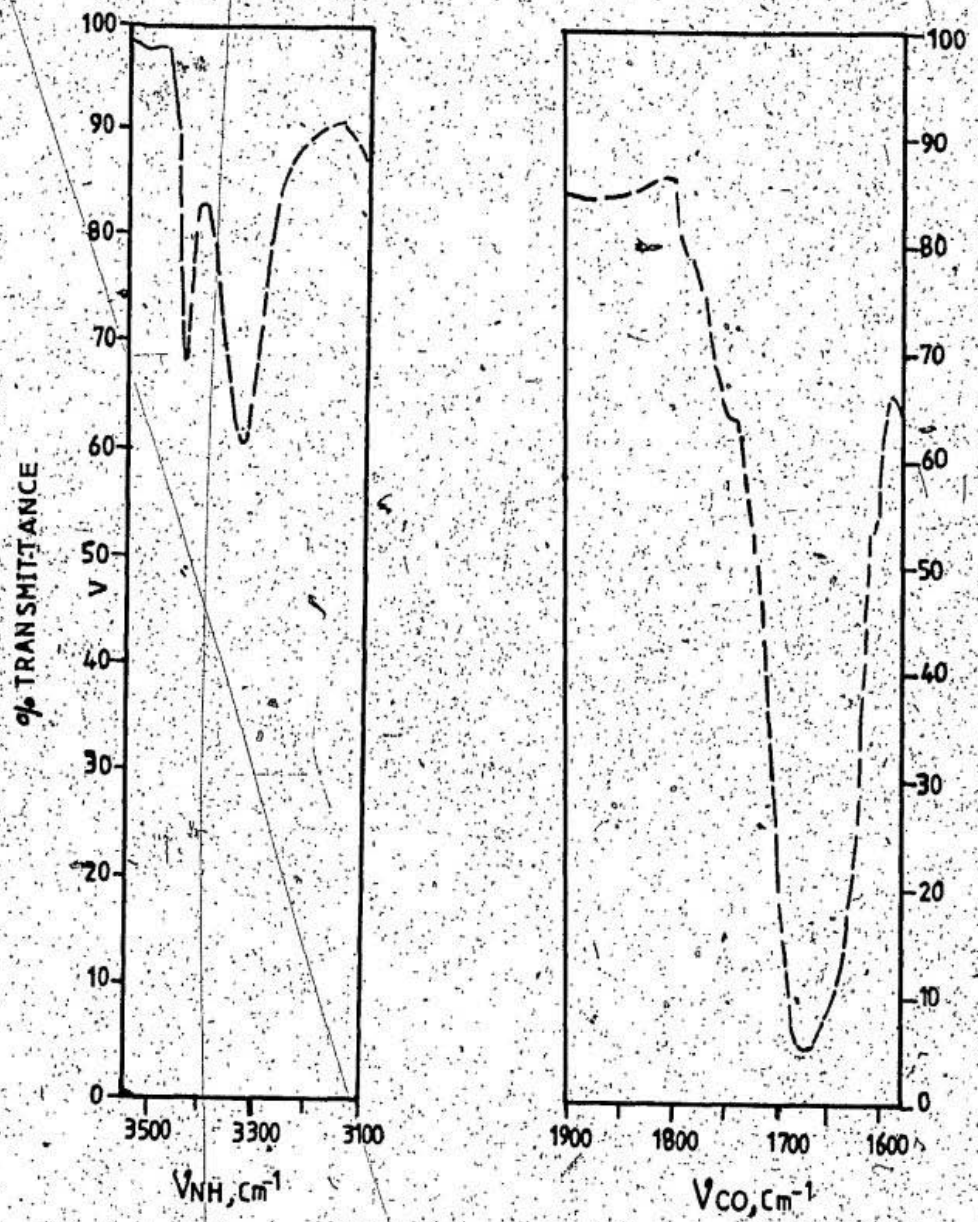


Figure 5-4: IR Spectra of t-Boc-Pro-Pro-Gly-$NHCH_3$ in $CHCl_3$

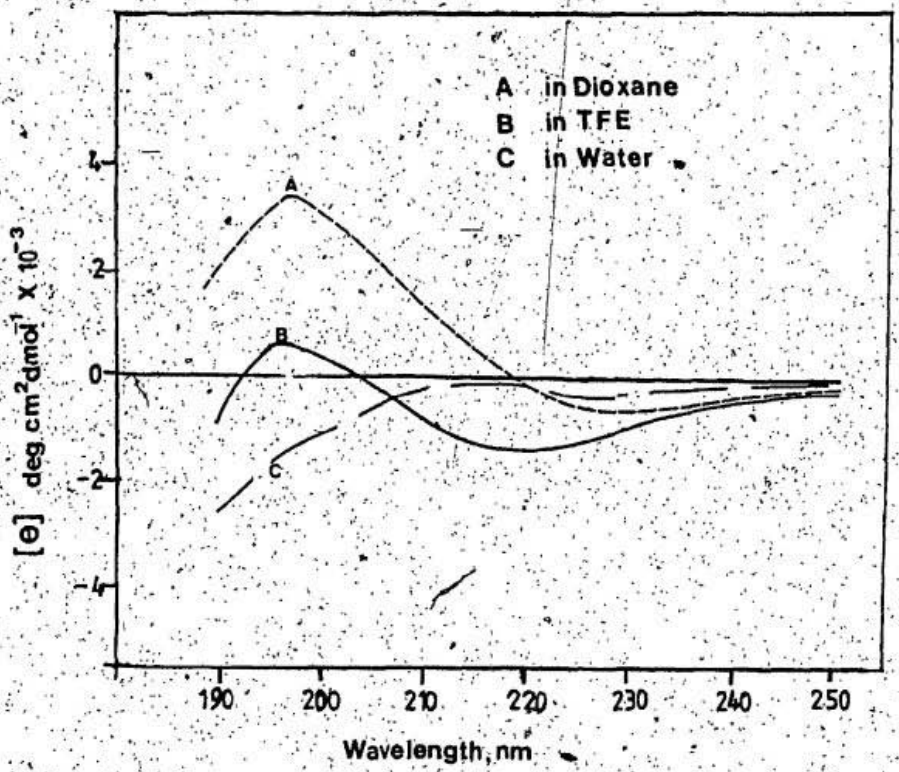


Figure 5-5: CD Spectra of t-Boc-Pro-Pro-Gly-$NHCH_3$ in dioxane (A), in TFE (B) and in water (C); concentration is 2 mg/ml. cell size 0.01 cm.

As discussed earlier in section 5.6, the spectra observed in TFE and dioxane are typical of class B spectra given usually by either type I or type II β-turns (Woody, 1974). The dioxane spectrum of the peptide also indicates the possibility of a 3 --> 1 intra-molecularly H-bonded conformer (Cann *et al.*, 1973; Madison and Schellman, 1970 a,b). In aqueous medium, however, the strong negative band seen at 190 nm indicates the predominance of an extended (PP-II like) conformer, relative to the folded conformer which is represented by the shallow negative trough around 225 nm. According to Madison and Kopple (1980) and Hollosi *et al.* (1985), in polar solvents like water, the chiral contribution of Pro residues dominates the CD spectrum of Pro-enriched peptides, resulting in the PP-II like extended conformers (with $\phi = -80^{\circ}$ and $\psi = 150^{\circ}$ i.e. trans-region). This conformation is characterized by a CD spectrum exhibiting a strong negative band around 200 nm.

Based on the above arguments, the spectral behaviour of Boc-PPG-$NHCH_3$ can be interpreted by assuming a conformational equilibrium between a PP-II like extended structure with a negative CD band around 200 nm and the β-turn (4 --> 1 H-bonded structure) exhibiting a negative band between 220-230 nm. This seems to be a reasonable assumption since other investigators have shown by spectroscopic and crystallographic studies that Pro-Pro sequences adopt an extended open conformation i.e. their ϕ and σ values are close to those of the PP-II structure (Benedetti *et al.*, 1983; Aubry *et al.*, 1985). On the other hand, Pro-Gly sequence adopts a "bent" structure (Tanaka *et al.*, 1977). Moreover, Pro-Gly sequence is among those found most often in β-turns (Lewis *et al.*, 1973; Ananthanarayanan *et al.*, 1984). The probability of β-turn formation for the Pro-

Gly sequence has been found to be very high based on the conformational energy calculations also (Stimson *et al.*, 1977; Lee *et al.*, 1984a).

The most supporting evidence for this structure can be obtained from X-ray studies on the almost identical tripeptide, t-Boc-Pro-Pro-Gly-NH_2 carried out by Tanaka *et al.*, (1979). These studies clearly demonstrated the presence of two well-defined conformations, an extended PP-II structure at the N-terminal half (t-Boc-Pro-Pro- segment) of the molecule and a β-turn at the C-terminal half, the C-terminal amide group acting as a H-donor to the CO group of Pro^1. The β-turn has been found to be of type I.

Based on this analysis, the structure of Boc-PPG-$NHCH_3$ can be schematically represented as shown in Figure 5-6. The crystal structure of Boc-PPG-NH_2 is also presented in Figure 5-7 for comparison.

5.6.3. t-Boc-Pro-Pro-Gly-Pro-OH

This tetrapeptide was selected to study the effect, on the conformation, of adding another Pro residue to the C-terminus of the tripeptide sequence Pro-Pro-Gly-, which is the minimum sequence to be hydroxylated by prolylhydroxylase.

The NH and CO stretching regions of t-Boc-Pro-Pro-Gly-Pro-OH (Boc-PPGP-OH) IR spectrum is shown in Figure 5-8. The data can be interpreted in an analogous manner to the tripeptide Boc-PPG-$NHCH_3$ presented earlier. The NH region shows "free" NH vibrations at 3400 cm^{-1} while the H-bonded NH band occurs at 3300 cm^{-1}. The relative intensities indicate the nearly equal dominance of the H-bonded and free conformers. The band above 3500 cm^{-1} is due to the

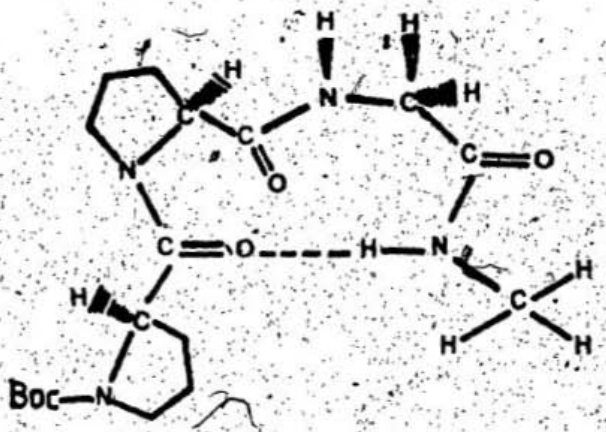


Figure 5-6: Possible Structure of t-Boc-Pro-Pro-Gly-$NHCH_3$ with a 4 --> 1 H-bonding

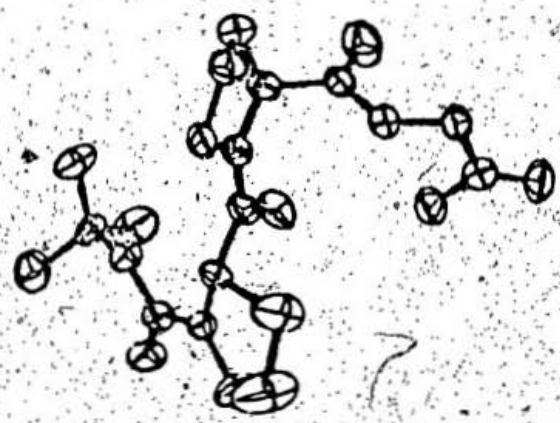


Figure 5-7: Crystal Structure of t-Boc-Pro-Pro-Gly-NH_2

From Tanaka *et al.*, 1979.

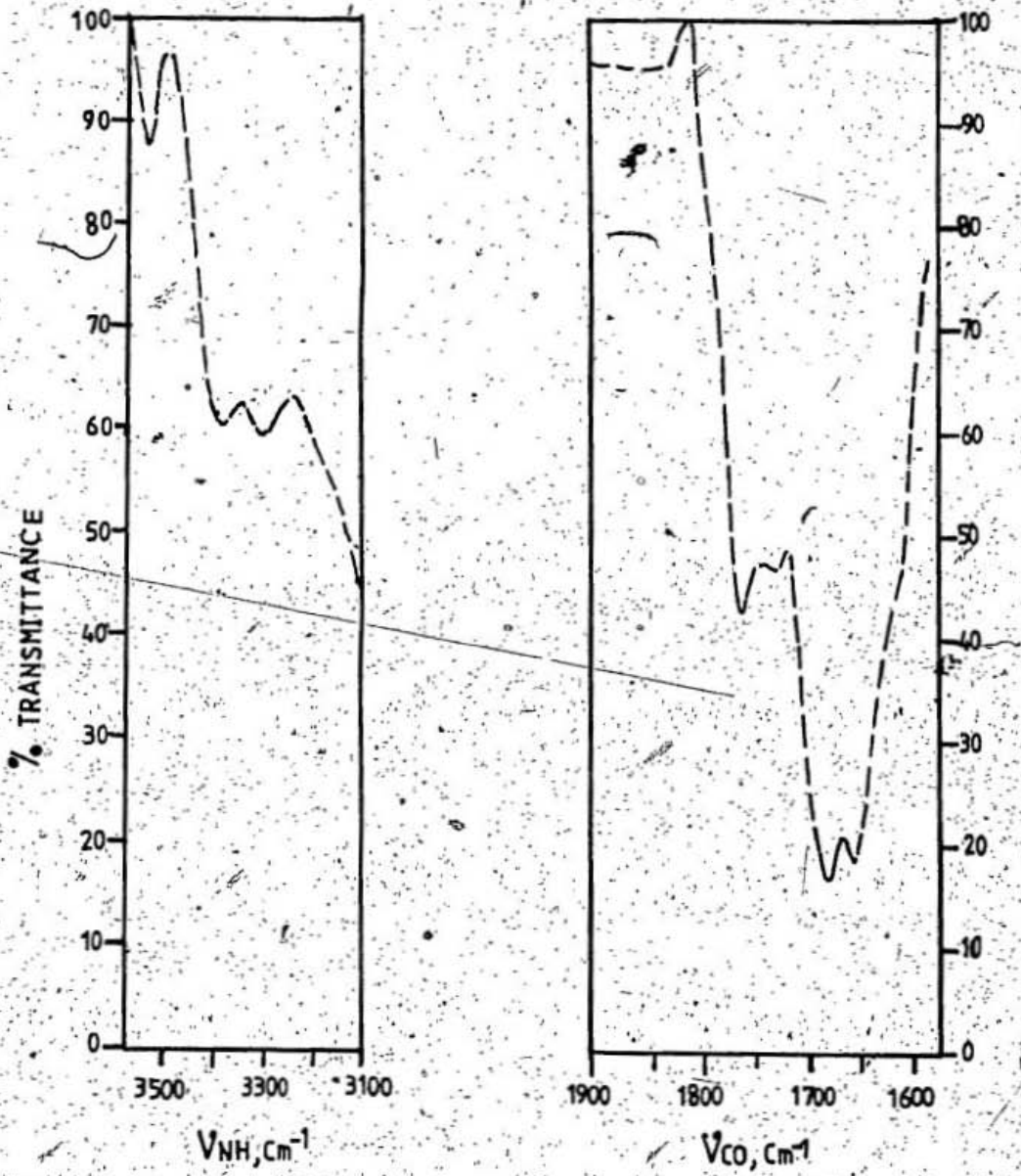


Figure 5-8: IR Spectra of t-Boc-Pro-Pro-Gly-Pro-OH in $CHCl_3$]

CO vibrations of the terminal COOH group. The CO stretching region shows a partly resolved broad band with a major shoulder at 1680 cm^{-1} and a minor one at 1650 cm^{-1}. . The presence of a 1670-1680 cm^{-1} carbonyl stretching band suggests the involvement of the peptide bond in H-bonding.(Hollosi *et al.*, 1985). The 1760 cm^{-1} band is due to the CO vibration of the terminal COOH group. The 1650 cm^{-1} shoulder can be ascribed to the carbonyl stretching vibrations of tertiary imide groups of prolyl residues in an extended form as was observed earlier by Lazarev *et al.* (1985) for PP-II (Isemura *et al.*, 1968) and other PP-II related polypeptides (Brahmachari *et al.*, 1978; Doyle *et al.*, 1975) and for Pro-containing linear tetrapeptides (Hollosi *et al.*, 1985).

Because the imino group of Pro residue following the Pro-Gly sequence cannot be a H-donor, a regular C_{10} structure with 4 --> 1 H-bonding may not be possible. The intra-molecular H-bonding indicated by the IR data would suggest other structures like the C_7 ring structures. This C_7 structure can be visualized to be formed between CO of Pro^1 and NH of Gly^3 residues. Crippen and Yang (1974) observed two frequencies at 3400 cm^{-1} and 3300 cm^{-1} for N-acetyl-Ala-N'Methylamide in which they proposed a similar C_7 conformation.

Turning now to the CD data, the CD spectra of Boc-PPGP-OH in TFE, water and dioxane are shown in Figure 5-9. In TFE, the spectrum displays a shallow negative shoulder around 230 nm and a strong negative band around 198 nm. In water, the spectrum shows a positive band around 228 nm and a strong negative band at 196 nm. In dioxane, the spectrum resembles a class C spectrum with two minima, one at 226 nm and another at 204 nm. The molar ellipticity at

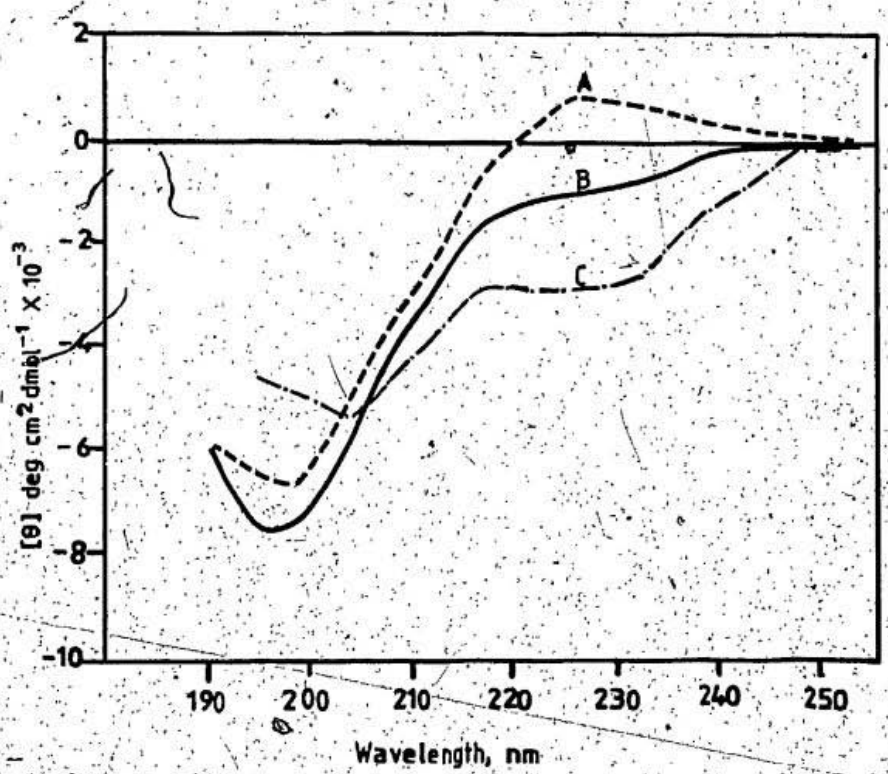


Figure 5-9: CD Spectra of t-Boc-Pro-Pro-Gly-Pro-OH in water (A), in TFE (B) and in dioxane (C); concentration = 2 mg/ml; cell size 0.01 cm.

204 nm is found to be more than that at 226 nm. The solvent dependence of the spectra indicate the conformational equilibrium between different structures.

The TFE spectrum of Boc-PPGP-OH is very similar to that observed in the case of CBZ-Gly-Pro-Leu-Gly-Ost and CBZ-Gly-Pro-Glu(OButyl)-Gly-Ost (st = stearyl) in acetonitrile by Hollosi *et al.* (1985) as shown in Figure 5-10. These spectra have been interpreted to be due to a conformational equilibrium between a PP-II like extended structure and a H-bonded type I β-turn. The negative band around 200 nm and shallow negative shoulder at 220 nm were assigned to PP-II and β-turn, respectively. In Boc-PPGP-OH, while the 198 nm band suggests a possible PP-II structure, the negative shoulder is found not at 220 nm as in the above case but at 230 nm. The negative band around 230 nm has been attributed to a C_7 or inverse γ-turn in N-acetyl-Pro-N'-methylamide (Madison and Schellmann, 1972; Cann *et al.*, 1973) and also in cyclo(Pro-Gly)$_3$ by Madison (1974). This structure contains a 3 --> 1 intra-molecular H-bond. In the present case, the IR data also support the possibility of a C_7 conformation with the 3 --> 1 intra-molecular H-bond between CO of Pro^1 and NH of Gly^3 residues.

The CD spectrum of the peptide in water (Figure 5-9) clearly indicates a shift in the conformational equilibrium in favour of the PP-II structure (at the expense of the γ-turn, whose intra-molecular H-bonding would be destabilized by the solvent, water). The PP-II structure is characterized by a positive CD band at 228 nm and a strong negative CD band around 196 nm. As a result of the larger Pro content, the PP-II structure in this tetrapeptide expresses itself more strongly (evidenced by the positive 228 nm band) than in the tripeptide,

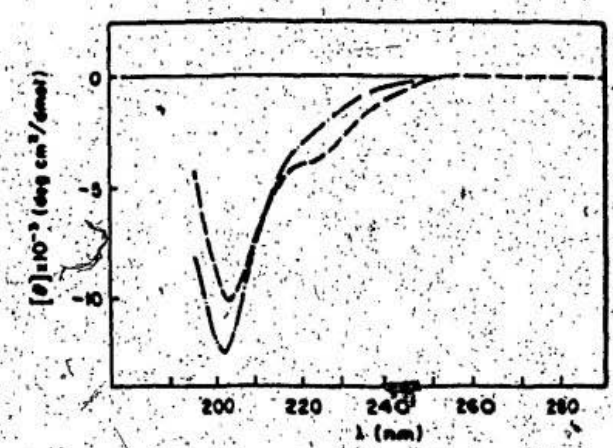


Figure 5-10: CD Spectra of CBZ-Gly-Pro-Leu-Gly-Ost in methanol (------), concentration = 0.6 mg/ml; in acetonitrile (- - -), concentration = 0,36 mg/ml.

From Hollosi *et al.*, 1985.

Boc-PPG-$NHCH_3$. Hence, a similar conformation can be attributed to the tetrapeptide as well, with an extended PP-II conformation followed by an inverse γ-turn which is stabilized by an intra-molecular 3 --> 1 H-bond between CO of Pro^1 and HN of Gly^3 residues. The proportions of the constituent structures, *viz.* PP-II and γ-turn, vary with the nature of the solvent.

5.6.4. t-Boc-Pro-Pro-Gly-Pro-$NHCH_3$

This peptide is homologous to the tetrapeptide discussed in the previous section and contains an additional peptide bond (between Pro-and N-methyl amide) at the C-terminus. The effect of this elongation at the C-terminus is examined in this section in terms of the spectral data.

The NH and CO stretching regions of the IR spectrum of Boc-Pro-Pro-Gly-Pro-$NHCH_3$ are shown in Figure 5-11. As seen from the figure, the free NH appears at 3450 cm^{-1} and 3400 cm^{-1} while H-bonded NH appears at 3340 cm^{-1}. The splitting of free NH, which was not discerned in the earlier peptides, might suggest the presence of cis- and trans-rotational isomers (Rao, 1963). The cis-trans isomerization is possible, in the present series of the peptides, at the X-Pro i.e. t-Boc-Pro (or Pro-Pro) bond. Pro-X is not prone to such isomerization. The band at 3400 cm^{-1} may also be indicative of the C_5 structures (Rao *et al.*, 1980) (which are formed with the trans-isomers only). However, the relative intensities of the free and bonded NH stretching bands suggest the predominance of H-bonded conformers. The CO stretching region also shows a broad, less resolved band at 1670 cm^{-1} with only a shoulder at 1650 cm^{-1}, indicating H-bonded peptide backbone. Since the Pro residue following the Pro-

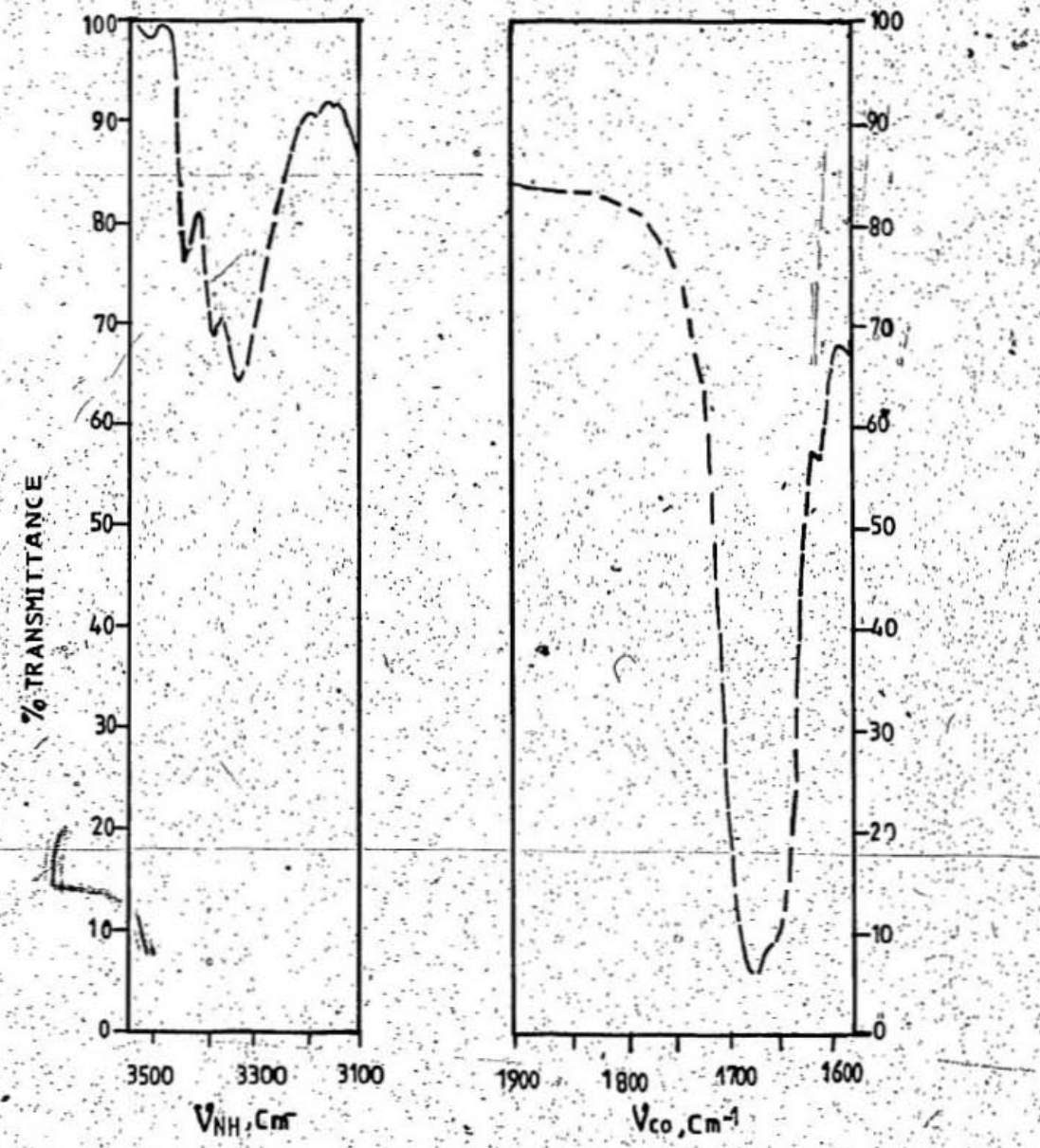


Figure 5-11: IR Spectra of t-Boc-Pro-Pro-Gly-Pro-$NHCH_3$ in $CHCl_3$

Gly sequence cannot participate in a regular C_{10} structure, the H-bonded species may denote a C_7 structure. C_5 structures are possible but not likely, by comparison with the tri- and tetrapeptide analogues. One can also see the absence of 1760 cm^{-1} band characteristic of the terminal COOH group since it is blocked in this case.

The CD spectra of Boc-PPGP-$NHCH_3$ in TFE, dioxane and water are shown in Figure-5-12. The spectrum in TFE shows a strong negative band around 198 nm with a faint negative shoulder around 230 nm. The spectrum in water is characterized by a positive band around 230 nm and strong negative band at 198 nm. The dioxane spectrum of the peptide shows an intense negative band at 226 nm and another less intense negative band around 208 nm. Based on a similar line of discussion as in the earlier cases, these features can be interpreted in terms of the different equilibrium proportions of the PP-II and γ-turn structures.

Conformational energy calculations have been carried out on N-acetyl-Pro-Gly-$NHCH_3$ to evaluate its potential for the formation of β-turns (Lee *et al.*, 1984a). It was shown that there was a high β-turn potential for the dipeptide, *viz.* 0.72, which increased further to 0.86 for the pentapeptide (N-acetyl-PPGP-$NHCH_3$). This increase was attributed to the stabilization of the β-turn by noncovalent interactions involving the first and last Pro residues. Incidentally, it is to be mentioned that both the first and the fourth Pro residues are reported in the so-called "F" state which favours the PP-II like extended structure. The structure of N-acetyl-PPGP-$NHCH_3$ arrived at by Lee *et al.*

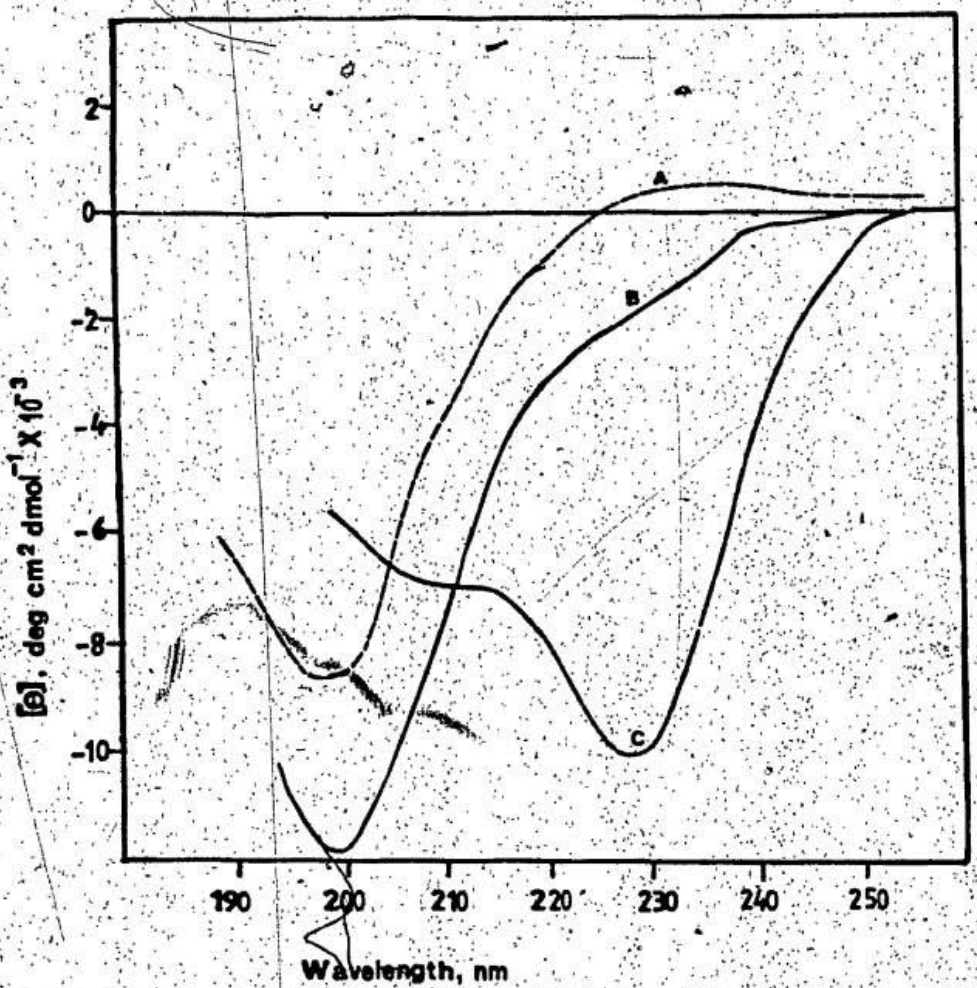


Figure 5-12: CD Spectra of t-Boc-Pro-Pro-Gly-Pro-$NHCH_3$ in water (A), in TFE (B) and in dioxane (C); concentration =2 mg/ml; cell size 0.01 cm.

(1984a), shown in Figure 5-13, can be regarded as having an "open" β-turn structure in addition to the PP-II structure at the N-terminal N-acetyl-Pro-Pro segment.

More recently, based on the conformational energy calculations, Bansal and Ananthanarayanan (1987) have indicated clearly that this "open" β-turn is, in fact, stabilized by a γ-turn involving the Pro^1 CO and Gly^3 NH groups. This is in agreement with the findings of Paul and Ramakrishnan (1985), who examined the minimum energy conformation of di- and tripeptides and observed the presence of an additional 3 --> 1 H-bond in some of the type I and type II β-turns in addition to the 4 --> 1 type H-bonding. These two types of H-bonds were found to occur in a bifurcated form as shown in Figure 5-14. This kind of H-bonding scheme seems to be possible also in Boc-PPGP-$NHCH_3$ and the consensus structure for this compound as derived from the experimental data and the above theoretical considerations, is shown in Figure 5-15.

5.6.5. t-Boc-Pro-Pro-Ala-Pro-OH, t-Boc-Pro-Pro-Gln-Pro-OMe and t-Boc-Pro-Pro-Gly-Pro-Pro-OH

These peptides were studied in view of the following considerations :

1. The former two peptides permit to examine the effect of introducing a side chain to the third, i.e. $(i+2)$nd residue on the formation of β-turn; while the study of the pentapeptide is expected to give information about the effect of elongation of the tetrapeptide by another Pro residue on its C-terminal side.

2. In some of the biologically important proteins like C1q (Reid and Porter

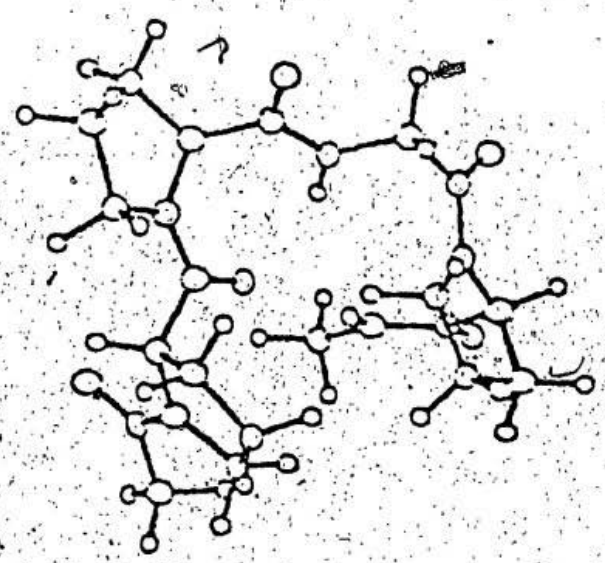


Figure 5-13: Minimum Energy Conformation of

N-acetyl-Pro-Pro-Gly-Pro-$NHCH_3$

From Lee *et al.*, 1984 a

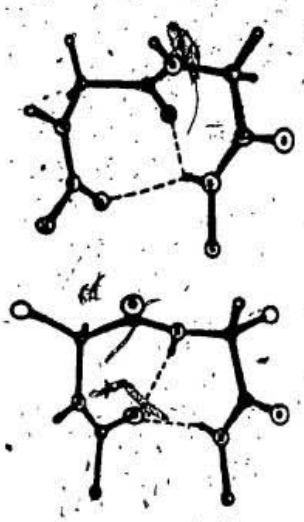


Figure 5-14: Bifurcated intra-molecular Hydrogen Bonds

From Paul and Ramakrishnan, 1985.

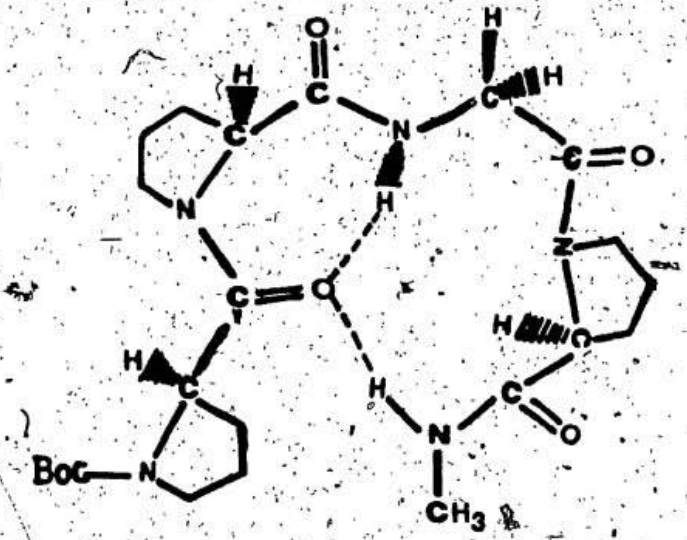


Figure 5-15: Possible Structure of t-Boc-Pro-Pro-Gly-Pro-$NHCH_3$

(1976) and Pro-rich salivary proteins (Bennick,1975) Pro-Ala and Pro-Gln sequences are found. It is of importance to know whether such sequences are recognized by prolylhydroxylase especially in the case of C1q where Hyp is also found.

The NH and CO regions in the IR spectrum of t-Boc-Pro-Pro-Gly-Pro-Pro-OH (Boc-PPGPP-OH) are shown in Figure 5-16. The NH region exhibits two bands, one at 3400 cm^{-1} and another at 3300 cm^{-1} whose relative intensities indicate the predominance of non-H-bonded or open conformers. The NH region in the IR spectrum of t-Boc-Pro-Pro-Ala-Pro-OH (Boc-PPAP-OH) is shown in Figure 5-17 which also exhibits two bands at 3410 and 3310 cm^{-1} respectively. In contrast to Boc-PPGPP-OH, the relative intensities of these bands in Boc-PPAP-OH indicate the predominance of the H-bonded conformers in this compound. On the other hand, the NH stretching region of t-Boc-Pro-Pro-Gln-Pro-OMe (Boc-PPQP-OMe) (shown in Figure 5-18) exhibits 4 bands; the 3480 cm^{-1} band is due to the free side chain primary amide of Gln residue. The 3400 cm^{-1} band is ascribable to free NH of the peptide amides not involved in H-bonding. The 3320 cm^{-1} and 3180 $^{-1}$ bands are due to H-bonded amide NH frequencies. The relative intensities of free and bonded NH stretching bands indicate the equal dominance of non-H-bonded and H-bonded conformers similar to the case of Boc-PPGP-OH (section 4.6.3).

Turning to the CO stretching region, Boc-PPGPP-OH in this region exhibits a broad, less resolved band at 1650 cm^{-1} (shown in Figure 5-16) with only a shoulder at 1680 cm^{-1} indicating that the predominant conformation of the

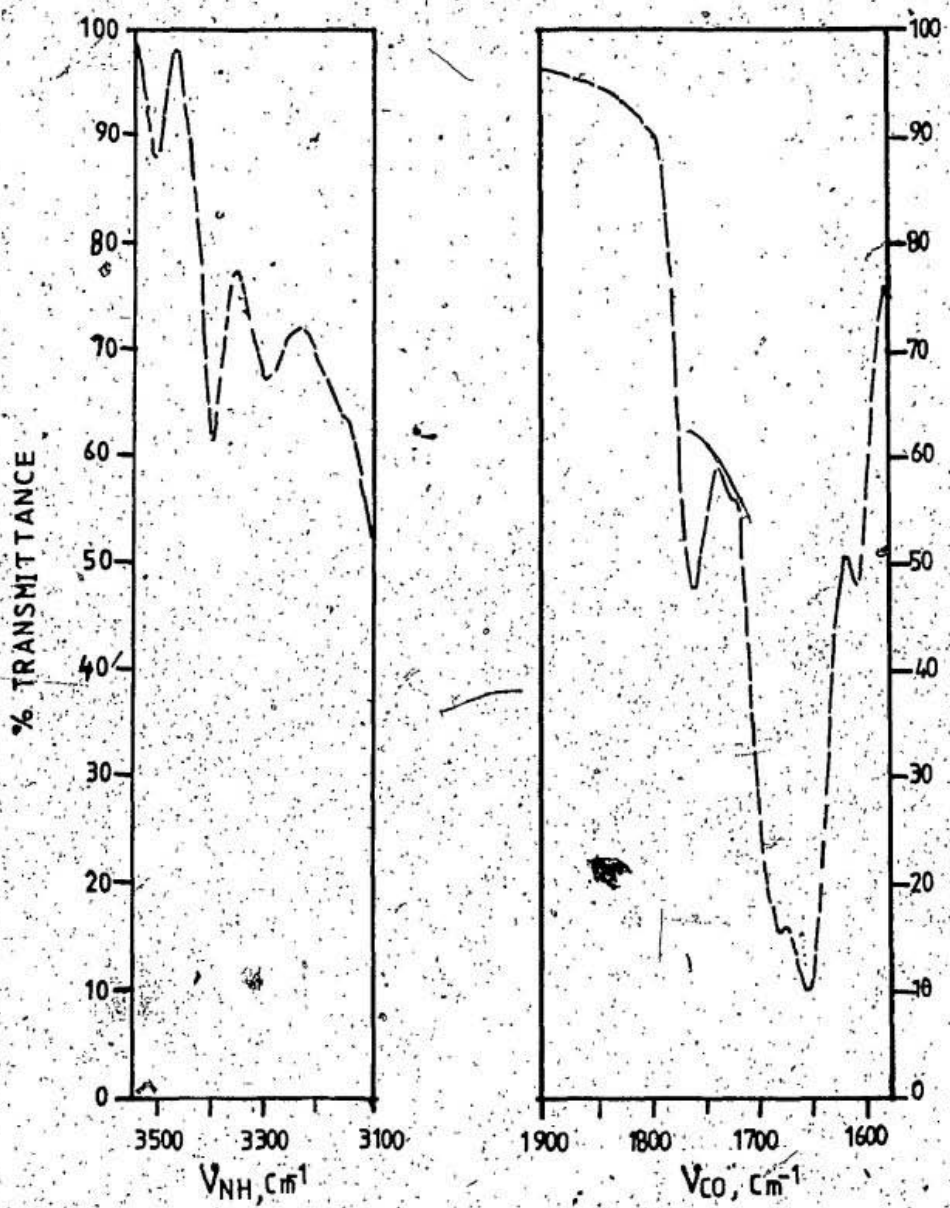


Figure 5-16: IR Spectra of t-Boc-Pro-Pro-Gly-Pro-Pro-OH in $CHCl_3$

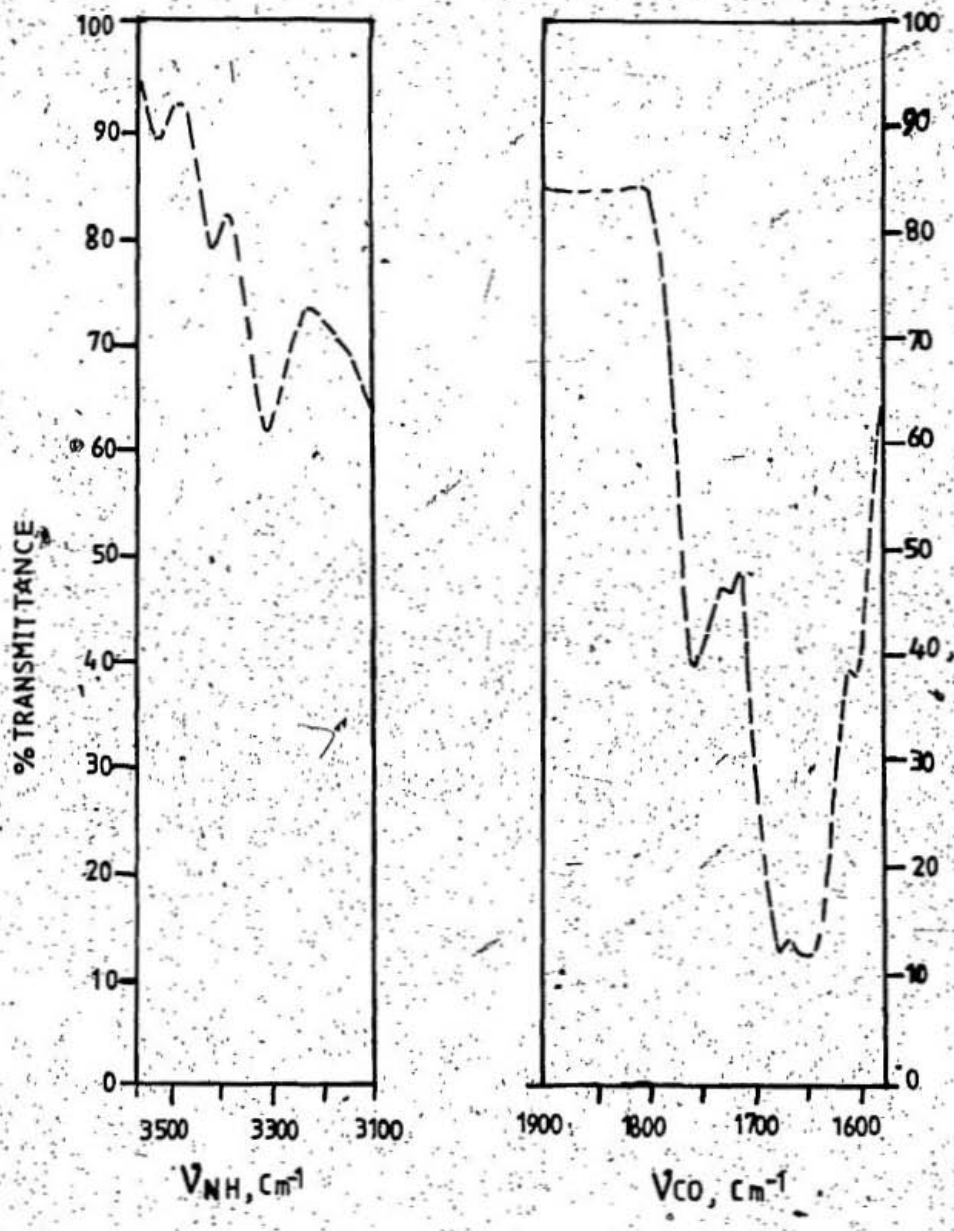


Figure 5-17: IR Spectra of t-Boc-Pro-Pro-Ala-Pro-OH in $CHCl_3$

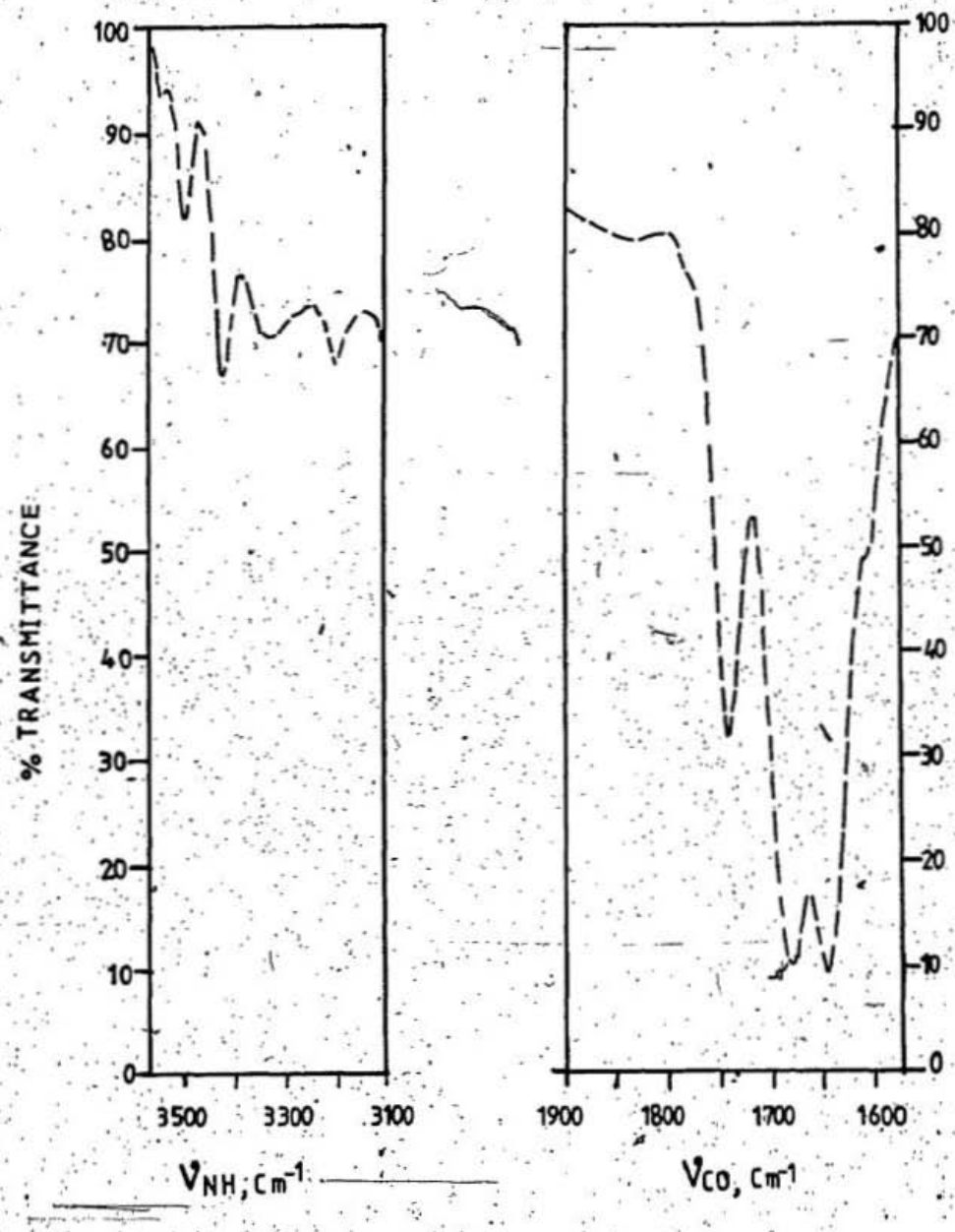


Figure 5-18: IR Spectra of t-Boc-Pro-Pro-Gln-Pro-OMe in $CHCl_3$

peptide backbone is the extended, open form. However, the smaller shoulder at 1680 cm^{-1} coupled with the smaller band with 3300 cm^{-1} represents a minor fraction of the conformers that can be stabilized by intra-molecular H-bonding. As discussed earlier in the case of Boc-PPGP-OH, the H-bonded conformation may be a C_7 structure between the CO of Pro^1 and HN of Gly^3 residues. In contrast to the above peptide, the CO stretching region of Boc-PPAP-OH shows a broad, partly resolved band bifurcated at 1680 and 1650 cm^{-1} (Figure 5-17) indicating the presence of both open and H-bonded conformers. In a very similar compound Piv-Pro-Pro-Ala-$NHCH_3$, Venkatachalapathi and Balaram (1979) observed a single prominent band at 3340 cm^{-1} which they ascribed to the presence of H-bonded 3_{10} helix. However, in Boc-PPAP-OH, the tertiary imide group of the Ala-Pro bond cannot act as a H-donor, unlike Piv-Pro-Pro-Ala-$NHCH_3$, where the NH of Ala methylamide is involved in intra-molecular H-bonding. Therefore, the observed intra-molecular H-bonding is likely to be between the CO of Pro^1 and NH of Ala^3, forming a C_7 structure.

The CO stretching region of Boc-PPQP-OMe (Figure 5-18) has a well-resolved band at 1740 cm^{-1} due to the ester carbonyl stretching. There are two partly resolved bands at 1680 cm^{-1} and 1645 cm^{-1} which indicate the presence of both H-bonded as well as extended conformations of the peptide bond as in the case of Boc-PPGP-OH and Boc-PPAP-OH. Hence, a C_7 conformation in equilibrium with an extended PP-II conformation can be visualized for this peptide also.

Figure 5-19 shows the CD spectrum of Boc-PPGPP-OH in TFE, dioxane

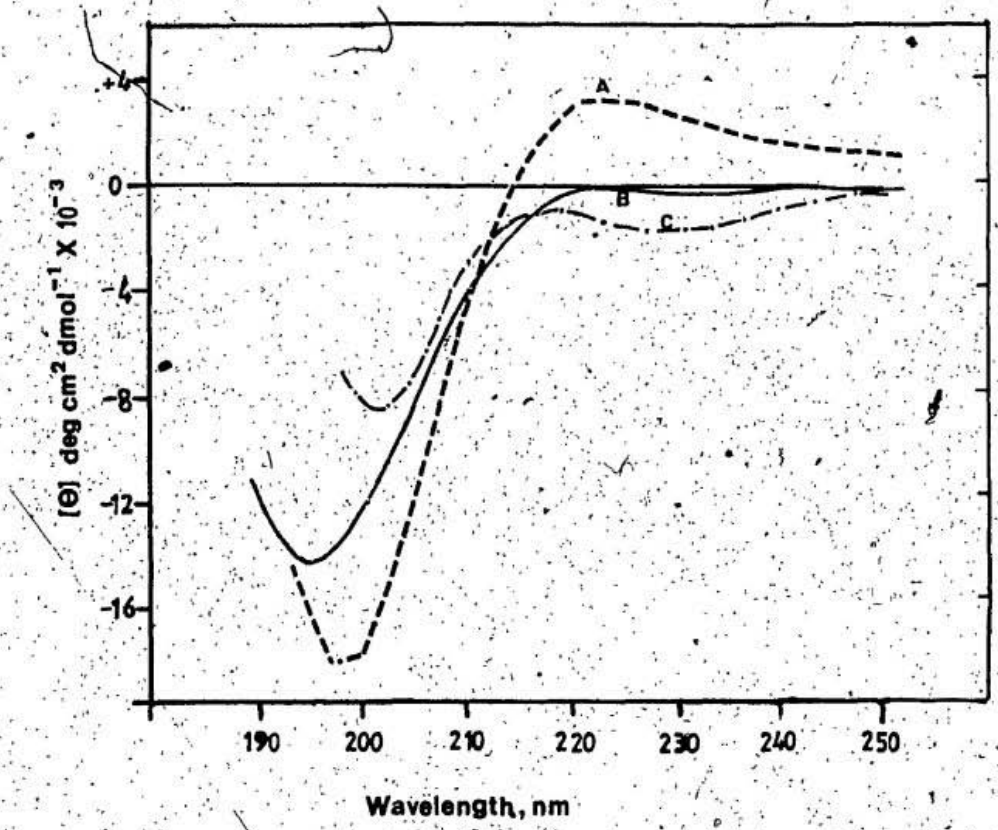


Figure 5-19: CD Spectra of t-Boc-Pro-Pro-Gly-Pro-Pro-OH

Spectra in water (A), in TFE (B) and in dioxane (C);

concentration = 2 mg/ml; cell size = 0.01 cm.

and water. In the TFE, two minima are observed, a shallow one at 230 nm and a strong band at 196 nm. In dioxane, a shallow negative band around 230 nm and a stronger negative band around 200 nm are seen. In water, a strong positive maximum is seen at 224 nm while the minimum is still seen at 196 nm. It may be noted that the CD spectra of this peptide in the three solvents are distinctly different from each other, as observed in all other peptides indicating the solvent dependence of conformational states. The presence of a shallow negative trough around 230 nm in TFE which becomes much stronger in dioxane is indicative of an intra-molecularly H-bonded γ-turn (Madison and Kopple, 1980). In aqueous solution, the spectrum is dominated by the chiral contribution of imino peptide groups resulting in a strong positive band around 224 nm and an intense minimum at 196 nm characterizing the spectrum of PP-II (Tiffany and Krimm, 1969 a,b). From their studies on Pro-containing peptides, Hollosi *et al.* (1985) concluded that the presence of even one Pro residue in a tetrapeptide sequence can significantly influence the conformation of the peptide, favouring PP-II like extended conformers especially in aqueous medium. Therefore, it is not surprising to see that Boc-PPGPP-OH with four Pro residues exhibits the PP-II like extended helix as the predominant conformation in an aqueous medium. Because of this, the contribution of imino groups [as manifested by the relative strength of the negative band around 200 nm (typical of PP-II structure) when compared to the shallow trough around 224 nm (characterizing H-bonded species)] can be observed even in H-bond favouring solvents. The IR data also seem to support this.

The CD spectra of Boc-PPAP-OH and Boc-PPQP-OMe in TFE, water and

dioxane are presented in Figure 5-20 and Figure 5-21, respectively. Both of these peptides seem to give spectra which are qualitatively similar to each other and also similar to those of Boc-PPGP-OH and Boc-PPGPP-OH. However, the band positions and intensities seem to be slightly different in each case.

In TFE, Boc-PPAP-OH exhibits a shallow negative shoulder around 230 nm and a very deep negative band around 200 nm. This spectrum is very similar to that of Boc-PPGP-OH and can be interpreted in an analogous manner. The TFE spectrum of Boc-PPQP-OMe also shows a shallow negative trough around 232 nm and a strong negative band around 202 nm. The negative bands around 200 nm are likely to be due to the extended PP-II structure while the shallow negative shoulder or trough around 230 nm can be attributed to the presence of intra-molecular H-bonded conformation. In water, both compounds exhibit strong negative bands around 200 nm and prominent positive bands at 226 nm which support the predominance of PP-II structure in an aqueous medium. In dioxane, two negative bands are seen at 230 nm and 208 nm in the case of Boc-PPAP-OH and at 236 nm and 206 nm in the Gln counterpart. These spectra also indicate a conformational equilibrium between two structures, a possible 3 --> 1 intra-molecularly H-bonded conformation and an extended PP-II like conformation. IR data discussed above also suggest the possibility of C_7 structures and extended structures. However, while Boc-PPAP-OH seems to exist more in a H-bonded conformation, Boc-PPQP-OMe seems to exist in rather open and extended conformation based on IR data.

In summary, all the above peptides exhibited similar spectral behaviour

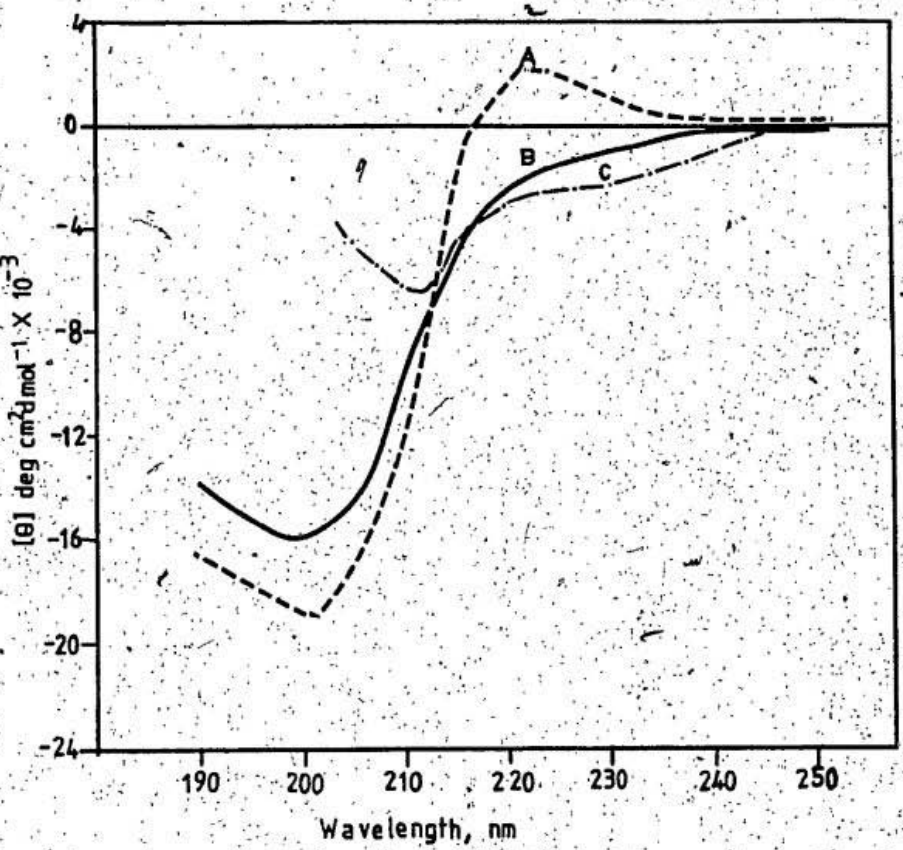


Figure 5-20: CD Spectra of t-Boc-Pro-Pro-Ala-Pro-OH

Spectra in water (A), in TFE (B) and in dioxane (C);

concentration = 2 mg/ml; cell size 0.01 cm.

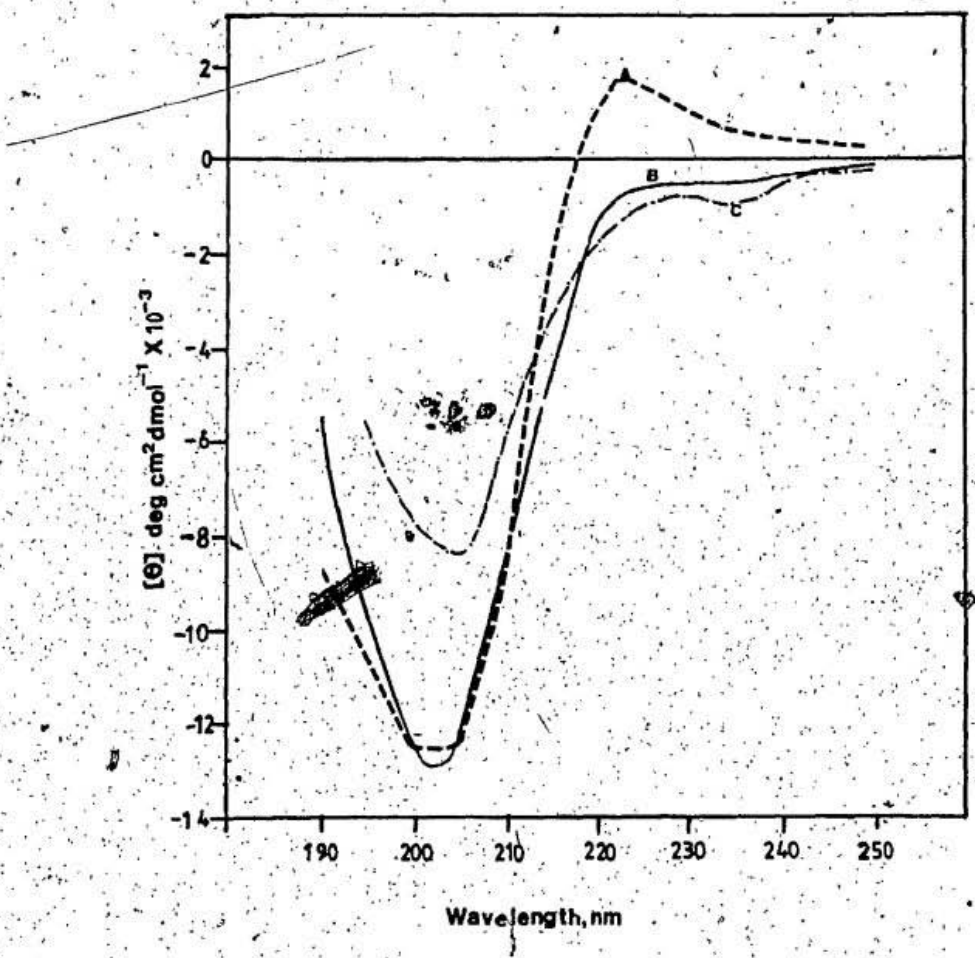


Figure 5-21: CD Spectra of t-Boc-Pro-Pro-Gln-Pro-OCH_3

Spectra in water (A), in TFE (B) and in dioxane (C);

concentration = 2 mg/ml; cell size 0.01 cm.

indicating similar conformational features when compared to each other and to Boc-PPGP-OH. This indicates that at least in some cases, Gly can be replaced by either Ala or Gln without drastic changes in the structure of the tetrapeptide sequences. The pentapeptide Boc-PPGPP-OH also exhibited essentially the same conformational features; however, due to the relative abundance of Pro residues in this peptide, the predominant conformation appeared to be an open, non-H-bonded, extended structure very similar to the PP-II left-handed helical conformation.

5.6.6. t-Boc-Val-Pro-Gly-Val-OH and t-Boc-Gly-Val-Pro-Gly-Val-OH

These peptides have been studied widely as elastin model peptides (Urry *et al.*, 1974a,b). Their use here allowed to examine the effect of introducing residues other than Pro on either side of the Pro-Gly sequence.

The IR spectra showing the NH and CO stretching regions of t-Boc-Val-Pro-Gly-Val-OH (Boc-VPGV-OH) and t-Boc-Gly-Val-Pro-Gly-Val-OH (Boc-GVPGV-OH) are shown in Figures 5-22 and 5-23, respectively.

The NH stretching region of Boc-VPGV-OH displays two major bands, one at 3440 cm^{-1} due to free NH vibrations and a bigger one at 3340 cm^{-1} representing the NH groups involved in H-bonding. The relative intensities indicate the predominance of H-bonded conformers. Boc-GVPGV-OH also exhibits similar features; the free NH vibrations appear at 3420 cm^{-1} as a small shoulder while the bigger 3340 cm^{-1} band is due to the major intra-molecularly H-bonded conformers. In the carbonyl stretching region, the small shoulder at

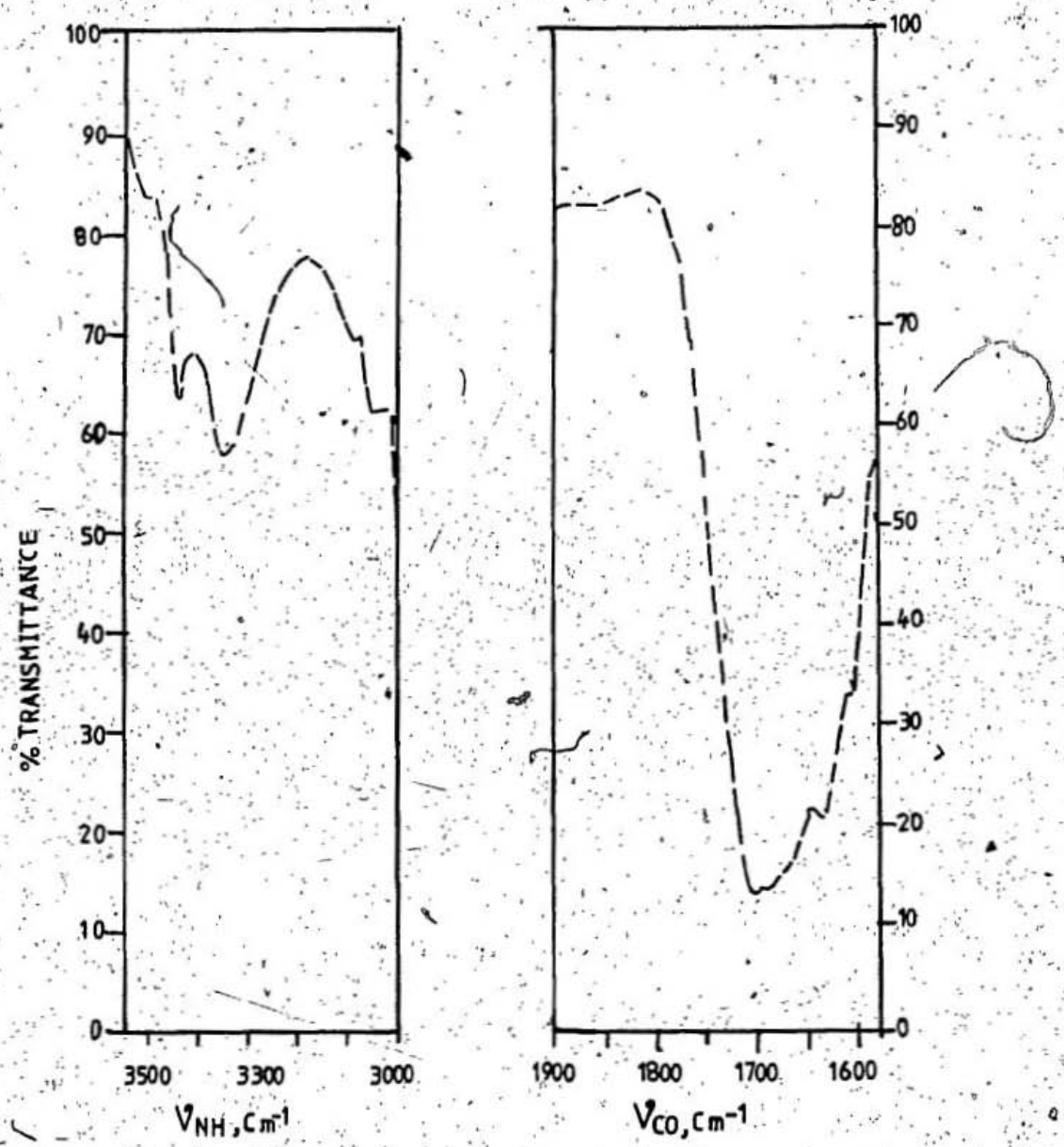


Figure 5-22: IR Spectra of t-Boc-Val-Pro-Gly-ValOH in $CHCl_3$

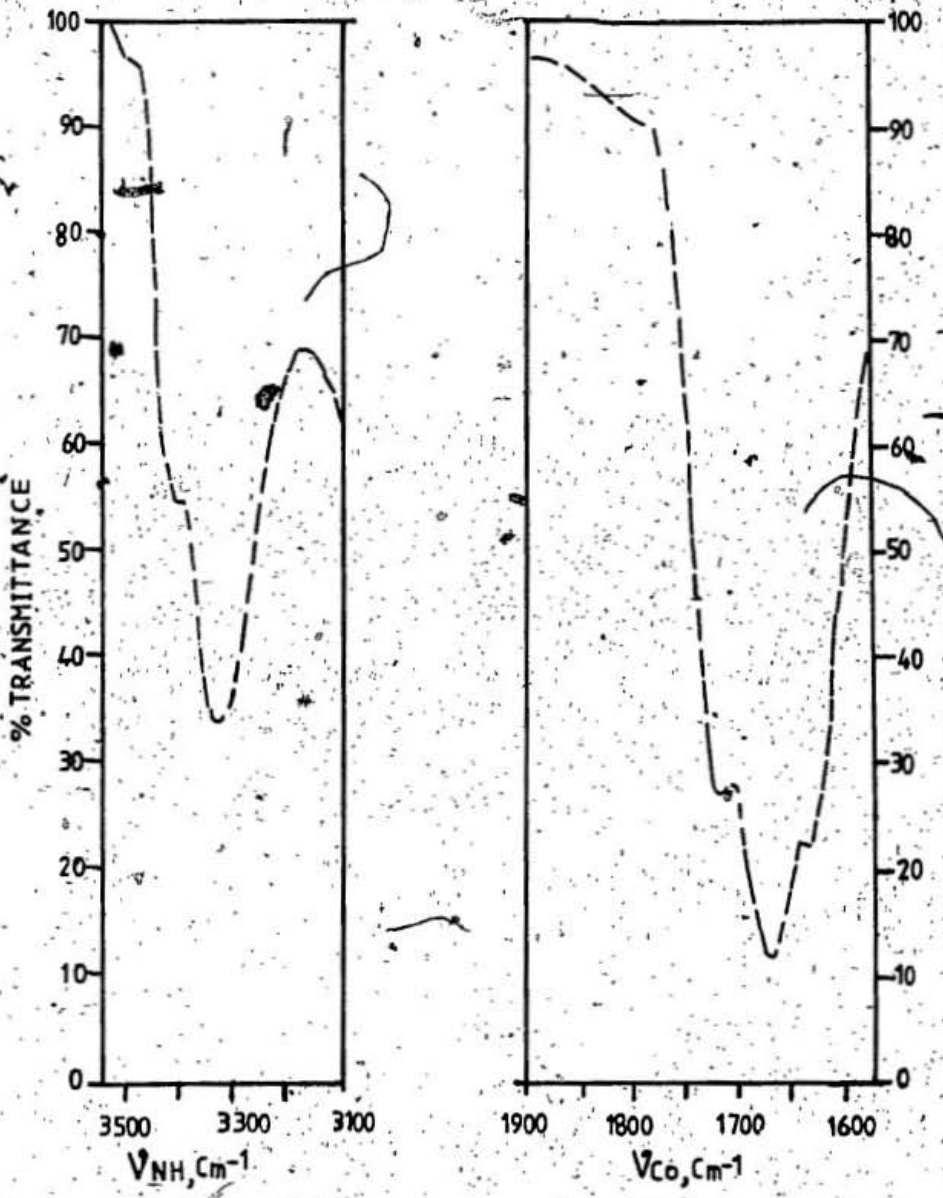


Figure 5-23: IR Spectra of t-Boc-Gly-Val-Pro-Gly-Val-OH in $CHCl_3$

1705-1720 cm^{-1} is due to the urethane carbonyl not involved in the H-bonding. The major band is at 1685 cm^{-1} in the case of Boc-VPGV-OH, while it is at 1670 cm^{-1} in the case of Boc-GVPGV-OH indicating, in both cases, the involvement of the peptide backbone in the intra-molecularly H-bonded conformations. The band positions of NH and CO stretching regions suggest 4 --> 1 type intra-molecular H-bonded conformers or C_{10} structures. The small 1630 cm^{-1} shoulder in both cases indicate the probable presence of a very minor fraction of inter-molecularly H-bonded species.

The CD spectra of Boc-VPGV-OH and Boc-GVPGV-OH in TFE and water are shown in Figure 5-24 and Figure 5-25, respectively. Both the peptides exhibit very similar CD spectral features indicating the similarity of their conformations. In TFE, the CD spectra of both the compounds are characteristic class B spectra (Woody, 1974) with a well-defined minimum around 224 nm and a maximum near 200 nm. However, the magnitudes of these extrema in the case of Boc-VPGV-OH are slightly higher indicating more ordered structure in this peptide. Class B spectra are known to be associated either with type I or type II β-turns. The peptide spectra in TFE are qualitatively very similar to those of cyclo(Val-Pro-Gly-Gly)$_3$ and HCO-(Val-Pro-Gly-Gly)$_n$-Val-OCH$_3$ and quantitatively intermediate to both of them (Urry *et al.*, 1974a). Based on extensive NMR studies, these peptides have been found to contain a type II β-turn with a 4 --> 1 intra-molecular H-bond between CO of Val^1 and NH of Gly^4 (Urry *et al.*, 1974b).

In water, the spectrum was dominated by two negative bands, one around

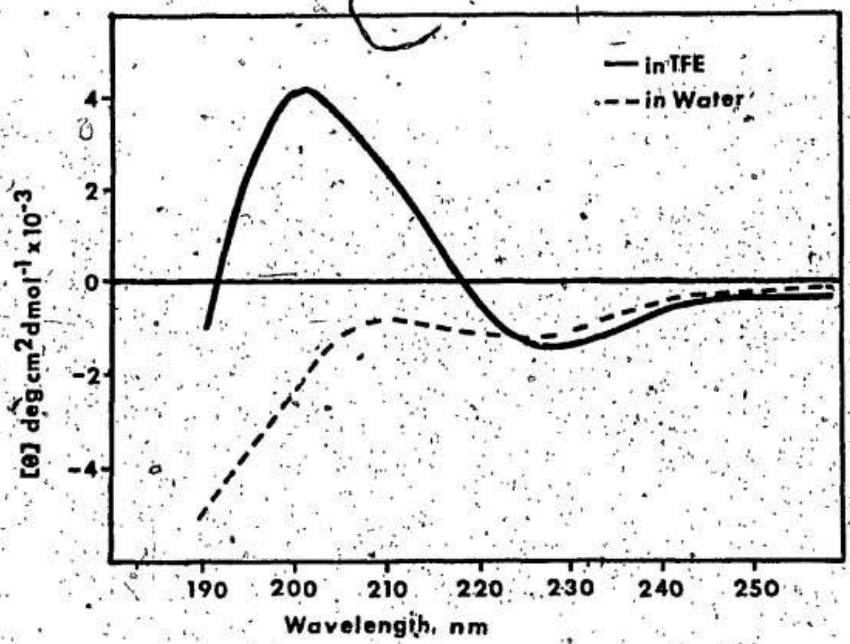


Figure 5-24: CD Spectra of t-Boc-Val-Pro-Gly-Val-OH

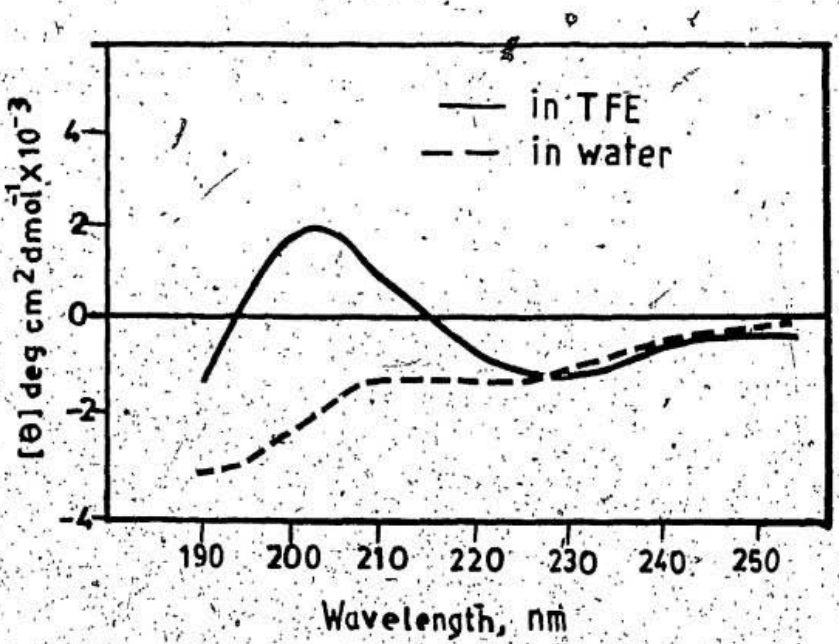


Figure 5-25: CD Spectra of t-Boc-Gly-Val-Pro-Gly-Val-OH

224 nm followed by a stronger negative band around 190 nm. This indicates the presence of two conformers in equilibrium with each other. The pronounced shallow trough around 225 nm suggests that the folded H-bonded conformation has not been completely disrupted in the aqueous medium indicating the strength of intra-molecular H-bond. The stronger negative band around 190 nm indicates the presence of either the unordered structure and/or an extended conformation. Unlike the case of earlier peptides, the nature of the extended conformation in these peptides seems to be rather unclear from their CD spectra.

X-ray studies by Yagi *et al.* (1983) on t-Boc-Val-Pro-Gly-Val-OH demonstrated an extended conformation in the t-Boc-Val-Pro- segment followed by a type II β-turn in the Pro-Gly segment of the molecule. The similar peptide, Boc-Val-Pro-Gly-Val-Gly-OH, has been found to take up an extended conformation in the solid state, although solution studies clearly suggested a type II β-turn existing in the molecule (Ayato *et al.*, 1980 and 1981; Urry *et al.*, 1983). This suggests the possible interconversion of folded and extended structures within the same peptide depending upon the environment. Yagi *et al.* (1983) have found that the extended structure in this case is similar to that of the antiparallel β-strand like structure. Based on these arguments, the CD spectral behaviour of Boc-Val-Pro-Gly-Val-OH and Boc-Gly-Val-Pro-Gly-Val-OH may be explained due to the presence of an extended, possibly β-sheet like structure followed by a β-turn. The possible structure for Boc-VPGV-OH is presented in Figure 5-26 and the crystal structure of the same compound as obtained from Yagi *et al.* (1983) is shown in Figure 5-27.

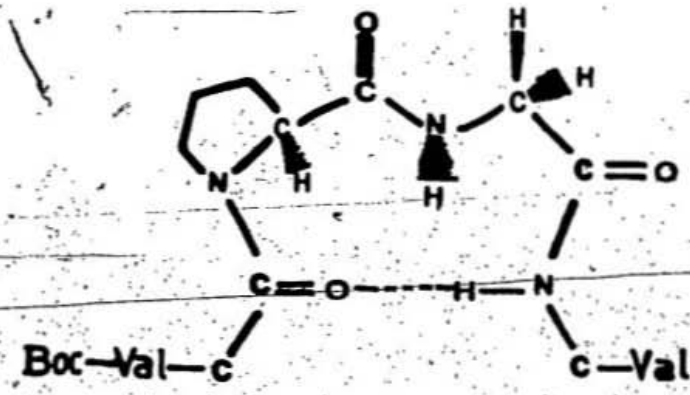


Figure 5-26: Possible Structure of t-Boc-Val-Pro-Gly-Val-OH with a 4 —> 1 intra-molecular H-bond

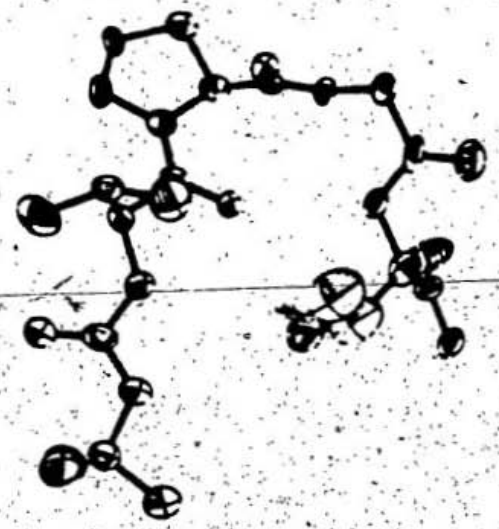


Figure 5-27: Crystal Structure of t-Boc-Val-Pro-Gly-Val-OH

From Yagi *et al.*, 1983.

5.7. Summary and Discussion

A general examination of these results indicate that :

1. Pro-containing peptides are capable of existing in an equilibrium mixture of multiple conformations, namely, β-turn, PP-II and unordered structures and

2. The nature of the solvent determines the different proportions of each of these conformations. In the organic solvent TFE, the folded, intra-molecularly H-bonded β-turn conformation is expresed more; on the other hand, in the aqueous medium, the extended conformation of PP-II becomes predominant. Thus, a given peptide would have larger proportions of the β-turn and smaller amounts of PP-II in TFE while in water the reverse would be the case.

However, in addition to these, the conformational population may include unordered conformations as well in equilibrium with the ordered structures. Therefore, it would be interesting to know the relative contributions of each of these conformers at equilibrium in a given solvent. An analysis to this effect is presented in the Appendix 1. A brief discussion about the results of this analysis is given below.

5.7.1. Determination of Proportions of Secondary Structures from CD Spectra of Peptides

The analysis was carried out by comparing the experimental CD spectrum of a given peptide in a given solvent with the computed spectra for various mixtures of the PP-II, β-turn and random coil structures as shown in Figure A-2 to A-9. In addition to the overall fit of the experimental spectrum with that of

the synthetic spectrum, the values at the wavelength maxima were given specific importance. Since the synthetic spectra were generated manually in steps of 10% variations of one or more conformations at each time, it is not surprising to see that the experimental spectra did not always match exactly with the computed spectra. In those instances, an intermediate position between two computed spectra was tentatively assigned to the experimental spectrum to obtain the percentages of the different conformers (f_{β}, $f_{PP\text{-}II}$ and f_{rc}). After this initial coarse curve-fitting, fine-tuning was carried out in order to obtain better curve-fitting. The results of this analysis are presented in Table 5-6.

Table 5-6: Conformational Analysis of Pro-containing Peptides

Peptide	Conformation in TFE	Conformation in water
	f_{β}: f_{PPII}: f_{rc}	f_{β}: f_{PPII}: f_{rc}
1.Boc-PPGNHCH$_3$	70:0:30	n.m
2.Boc-PPGP-OH	45:0:55	35:45:20
3.Boc-PPGPNHCH$_3$	30:0:70	40:40:20
4.Boc-PPGPP-OH	50:25:25	10:85:5
5.Boc-PPAP-OH	40:20:40	20:60:20
6.Boc-PPQP-OMe	55:25:20	20:60:20
7.Boc-VPGV-OH	85:0:15	65:0:35
8.Boc-GVPGV-OH	77:0:23	60:0:40
9.(Pro-Pro-Gly)$_5$	40:20:40	35:25:40

n.m Not matched; i.e. poor fit between the experimental and the computed spectra

5.7.2. Discussion of the Results from CD Data Analysis

5.7.2.1. Conformational Analysis of (Pro-Pro-Gly)$_5$

The experimental and computed spectra for this standard peptide substrate are presented in Figures A-25 and A-26. The analysis indicates about 35% β-turn, 25% PP-II and 40% random-coil structure present in aqueous medium. In TFE, about 40% β-turn, 20% PP-II and 40% random-coil structures were indicated. The analysis thus indicates that in the standard substrate of prolylhydroxylase, folded β-turn and extended PP-II-like conformations are present in equilibrium with flexible unordered or random-coil structure (both in H-bond promoting TFE as well as polar aqueous medium). One can see a slight increase in the PP-II-like extended structures in the aqueous medium when compared to TFE.

5.7.2.2. Conformational Analysis of t-Boc-Pro-Pro-Gly-NHCH$_3$, t-Boc-Pro-Pro-Gly-Pro-OH and t-Boc-Pro-Pro-Gly-Pro-NHCH$_3$

In the case of these peptides, the analysis indicates that in TFE, only the β-turn conformation is present in equilibrium with the unordered structure; the PP-II structure being insignificant (see Figures A.11 to A.15). Of these peptides, t-Boc-Pro-Pro-Gly-NHCH$_3$ (Boc-PPG-NHCH$_3$) has the highest proportion of β-turn conformation (70%). On going from TFE to aqueous medium, although there seemed to be a definite shift in the conformational equilibrium, the Boc-PPG-NHCH$_3$ spectrum in aqueous medium could not be matched well with any of the computed spectra within reasonable ($<$ 10%) standard deviation. Data on this and similar cases were therefore considered to be not very useful.

In the cases of Boc-PPGP-OH and Boc-PPGP-NHCH$_3$, which exhibit a

positive CD band around 225 nm (in aqueous medium), only the values at this wavelength region could be matched well. The values at the lower wavelength negative band are quite different probably due to the very pronounced effect of chain-length on this band as discussed in the Appendix. The values shown for f_{PP-II} in these two peptides were therefore, obtained from the analysis of the 225 nm band only.

5.7.2.3. Conformational Analysis of Boc-PPGRP-OH, Boc-PPAP-OH and Boc-PPQP-OMe

The experimental and computed spectra of these peptides are shown in Figures A-16 to A-20. The analysis indicates the similarity of conformational equilibria among these peptides. Unlike the first set of peptides mentioned above, these peptides exhibit the extended PP-II conformers even in TFE, to the extent of 20-25% of the total conformers. However, in TFE, β-turn conformation is still the dominant conformation making up as much as 50% of the total conformers. The change of the solvent from TFE to aqueous phase results in the reorganization of the conformers so that the most dominant one is the PP-II like extended conformation, constituting as much as 65-85% of the total conformers. The pentapeptide Boc-PPGPP-OH exhibits the highest proportion of PP-II conformers (85%) which is not unexpected, since it has the highest proline content. In the case of these peptides also, the spectra in water could be matched well only near the positive band. The values at the lower wavelength negative band of the computed spectra are much higher than those of the experimental spectra.

An examination of the computed data obtained for the above peptides (Table 5-6) indicates that the latter two favour the extended conformers more when compared to Boc-PPGP-OH. This may be due to the steric limitations imposed by the larger side chains of Ala and Gln when compared to Gly.

5.7.2.4. Conformational Analysis of Boc-VPGV-OH and Boc-GVPGV-OH

The experimental and computed spectra for these compounds are presented in Figures A-21 to A-24. The analyses indicate that the major conformers (75-85%) exist as folded β-turn structures in equilibrium with unordered conformers in TFE. Even in aqueous medium, the presence of the PP-II structure in these peptides was not indicated by the curve-fitting analysis. Thus, this analysis also indicates that the extended structure present in these peptides in aqueous medium is different from the PP-II like structure. However, since only β-turn, PP-II and random-coil structures are considered as the reference spectra, the present analysis would not be able to give information on the nature of extended structure present in these peptides. Consideration of the CD spectra of other extended structures like β-sheet structure may improve the conformational analysis on these two peptides. In addition to the above manual analysis, the analysis was also attempted using a computer program. The values given by the computer were found to be very similar to those obtained by manual analysis in some cases; however, in other cases, the curve-fitting by the computer was totally unacceptable in view of the original spectral characteristics of these peptides. The reason for this is not understood and attempts to improve the programme are underway.

Chapter 6

Interaction of Pro-containing Oligopeptides with Prolylhydroxylase

6.1. Introduction

As discussed in Chapter 1, earlier studies have indicated that prolylhydroxylase specifically hydroxylates proline residues in X-Pro-Gly sequences of its polypeptide substrates (Prockop *et al.*, 1976; Bornstein and Traub, 1979). The nature of the adjoining X-residue seems to have an effect on the extent of hydroxylation of the proline residue in the X-Pro-Gly segment (Kivirikko *et al.*, 1968; Prockop *et al.*, 1976 and Rapaka *et al.*, 1978).

Based on both theoretical and experimental considerations, Brahmachari and Ananthanarayanan (1979) proposed that the β-turn conformation adopted by the Pro-Gly segments in nascent procollagen molecules is specifically recognized by prolylhydroxylase and this is the conformational basis for the enzyme's specificity towards X-Pro-Gly segments as opposed to the Gly-X-Pro segments. These authors also proposed that the extent of enzymatic hydroxylation of proline residues, which is known to be influenced by the nature of the adjoining residue (Prockop *et al.*, 1976), may be related to the extent of stabilization of β-turn in a given tripeptide sequence. This hypothesis thus offers an explanation, in

conformational terms, for the specific recognition and hydroxylation of X-Pro-Gly segments in nascent procollagen.

More recently, Chopra and Ananthanarayanan (1982) provided experimental data which support both these postulates. The following Pro-Gly-containing peptides were synthesized : t-Boc-Pro-Gly-Ala-OH, t-Boc-Pro-Gly-Val-OH, t-Boc-Pro-DAla-Ala-OH and t-BOc-Gly-Val-Pro-Gly-Val-OH. CD, IR and NMR spectral data showed that these peptides contained the β-turn conformation (Ananthanarayanan and Shyamasundar, 1981; Brahmachari *et al.*, 1982). All of these peptides were found to inhibit the enzyme effectively. Interestingly, the pentapeptide was hydroxylated itself and in addition, it could inhibit the hydroxylation of the standard substrate. These results lent support to the β-turn hypothesis by showing that this conformation present in Pro-Gly segments of the substrate is recognized at the enzyme's active site.

However, closer examination of the above data reveals that the tripeptide sequence containing the β-turn is a necessary but not sufficient prerequisite for proline hydroxylation. Proline hydroxylation also seems to depend on the presence of the additional residues. This is illustrated by considering the peptides t-Boc-Pro-Gly-Val-OH and t-Boc-Gly-Val-Pro-Gly-Val-OH both containing the β-turn conformation (Chapter 5). The former is not a substrate and only can act as an inhibitor while the latter can not only inhibit the hydroxylation of the standard substrate, but also can undergo hydroxylation by prolylhydroxylase. Therefore, further studies are necessary to understand the contribution of the additional residues, in enhancing the interaction between the minimum β-turn

sequences and the enzyme. This may be related to either the so-called "chainlength effect" or the stabilization of the β-turn by the additional residues. In what follows, a detailed investigation of the interaction of prolylhydroxylase with peptides of varying lengths, composition, sequence and conformation will be presented, starting with di- and tripeptides.

6.2. Studies on Di- and Tripeptides

Studies on several dipeptides and tripeptides were carried out to understand their interaction with prolylhydroxylase. Table 6-1 lists the di- and tripeptides which acted as inhibitors of prolylhydroxylase with respect to the standard substrate, namely, $(Pro\text{-}Pro\text{-}Gly)_5$. The conformation of these peptides as known from the literature is also included.

From the above data, it becomes clear that not only the β-turn-containing tripeptides like t-Boc-Pro-Gly-Val-OH but also the dipeptides *viz.* t-Boc-Pro-Gly-OH, t-Boc-Pro-DAla-OH and t-Boc-Gly-Pro-OH and the tripeptide t-Boc-Gly-Pro-Pro-OH (none of which can sustain the β-turn conformation) acted as inhibitors to the enzyme with respect to the regular substrate. It is to be noted that some of these dipeptides adopt a "bent" structure, while others take up a rigid extended structure, both of which are different from the typical β-turn conformation (see section 6.7).

Thus, the initial inhibition data seemed to suggest that prolylhydroxylase might recognize non-β-turn conformations as well. Even when the β-turn conformation prevails as in Boc-Pro-Gly-X-OH tripeptides, one does not observe

Table 6-1: Inhibition of Prolylhydroxylase by Pro-containing Di- and Tripeptides

Peptide#	%Inhibition*	Ref.	structure	Ref.
Boc-P-G-A	90± 5	1	β-turn	2
Boc-P-G-V	90± 5	1	β-turn	2
Boc-P-DA-A	70± 3	1	β-turn	2
Boc-G-P-P	60± 6	3	Extended	3
Boc-P-G	40± 10	4	'Bent'	3
Boc-P-DA	50± 5	4	'Bent'	6
Boc-G-P	35± 15	3	'Rigid'	7
Boc-P-P	40± 10	3	'Rigid'	8

Peptide composition is denoted in one letter code for amino acids

Inhibitor concentration was 10 mM in the case of tripeptides and 20 mM for dipeptides

* Expressed with respect to $(Pro\text{-}Pro\text{-}Gly)_5$ as the substrate

"bent" conformation denotes an 'γ-shaped' structure

Ref. Reference

1. Chopra and Ananthanarayanan, 1982

2. Brahmachari and Ananthanarayanan, 1979

3. This thesis

4. Chopra and Ananthanarayanan, unpublished

5. Benedetti *et al.*, 1977

6. Ananthanarayanan and Cameron, Int. J. Pept. Prot. Res. (in press)

7. Tanaka *et al.*, 1977

8. Aubry *et al.*, 1985

hydroxylation. It appears, therefore, that the additional residues play a definite and important role in making the inhibitor into a substrate. Further, it is known that PP-II in the characteristic extended conformation and native triple-helical collagen are effective inhibitors of prolylhydroxylase although neither of them contains β-turns in their structures. Therefore, the conformational criteria for proline hydroxylation appear intriguing and this calls for additional experimental data in order to delineate the intricate possibilities. A revision of the earlier β-turn model was proposed by Ananthanarayanan (1984), which involved a combination of both the PP-II and β-turn conformation. This model is shown in Figure 6-1. According to this model, the active site of prolylhydroxylase consists of a binding site which requires the PP-II conformation (the "PP-II arm") and a catalytic site where the β-turn conformation is the required structure. The present studies were carried out to specifically test this new model.

6.3. Further Studies with Pro-containing Oligopeptides

Following the same approach used earlier, a series of Pro-containing oligopeptides (n = 3-5) were designed such that they would have additional residues besides those minimally needed to form the β-turn conformers. These peptides are listed in Table 6-2.

The choice of these peptides was based on the following considerations:

1. Proline usually occurs in the X position of Gly-X-Y repeating sequences of nascent collagen where the Y position can be occupied by any residue including Pro or Hyp, but not Gly. Therefore, the hydroxylatable Pro-Gly sequences

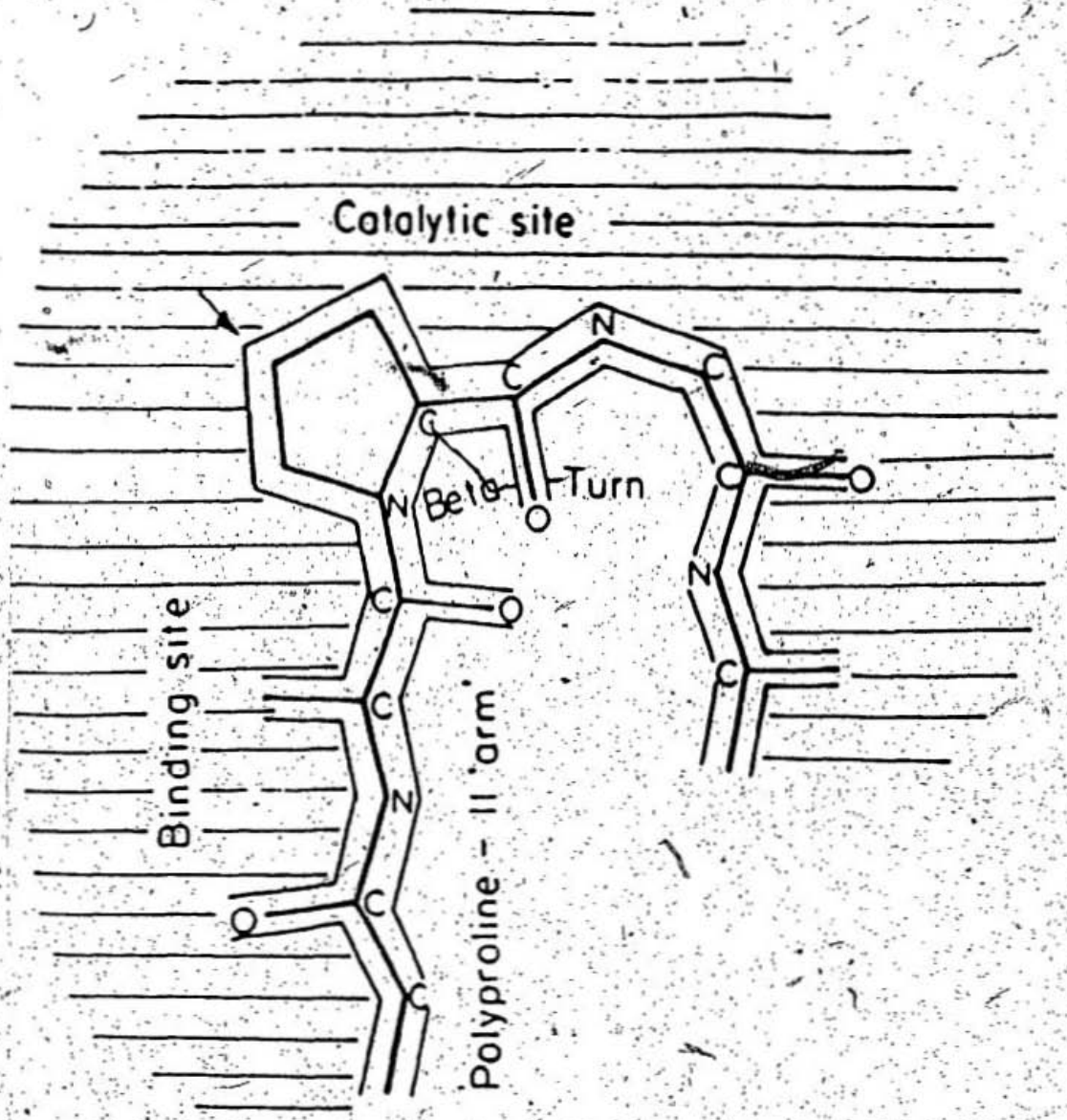


Figure 6-1: Schematic Representation of the Conformational Model for Prolylhydroxylase Substrates

Table 6-2: Pro-containing Oligopeptides Selected for Interaction with Prolylhydroxylase

Peptide	Abbreviation
1. t-Boc-Pro-Gly-Val-OH	Boc-PGV-OH
2. t-Boc-Pro-Gly-Leu-OH	Boc-PGL-OH
3. t-Boc-Pro-Pro-Gly-$NHCH_3$	Boc-PPG-$NHCH_3$
4. t-Boc-Pro-Pro-Gly-Pro-OH	Boc-PPGP-OH
5. t-Boc-Pro-Pro-Gly-Pro-$NHCH_3$	Boc-PPGP$NHCH_3$
6. t-Boc-Pro-Pro-Gly-Pro-Pro-OH	Boc-PPGPP-OH
7. t-Boc-Pro-Pro-Ala-Pro-OH	Boc-PPAP-OH
8. t-Boc-Pro-Pro-Gln-Pro-OH	Boc-PPQP-OMe
9. t-Boc-Val-Pro-Gly-Val-OH	Boc-VPGV-OH
10. t-Boc-Gly-Val-Pro-Gly-Val-OH	Boc-GVPGV-OH

flanked on either side by Pro residues would be good candidates for study in terms of their conformation and interaction with the enzyme. This is especially important for understanding why the "standard" polypeptide, namely, $(Pro\text{-}Pro\text{-}Gly)_n$ (where n = 5 or 10 residues) gets hydroxylated very well by the enzyme. No data, however, were available (until recently; see later) to show the presence of β-turns (at the Pro-Pro-Gly-Pro repeating sequences) in these peptides before they are hydroxylated.

2. Starting from the tripeptide level, increasing the number of additional residues on either side of Pro-Gly sequence, as in the case of peptides 1-6 shown in Table 6-2, can give information not only regarding the chainlength effect but also about the effect of these residues on the β-turn conformation formed by the Pro-Gly sequences.

3. Boc-PPAP-OH and Boc-PPQP-OMe were chosen to test whether the Pro-Gly sequence is indeed a near-absolute requirement for prolylhydroxylase as claimed by others (Prockop *et al.*, 1976). These peptide sequences allowed one to examine the effect of replacement of Gly by other amino acid residues such as Ala or Gln. The study of these peptide sequences is also of importance in the light of non-collagenous, Hyp-containing biological proteins like C1q, where Pro-Ala sequences are found in addition to Pro-Gly sequences (Reid and Porter, 1976). Also, a proline-rich human salivary protein which undergoes hydroxylation by prolylhydroxylase (data not shown) is found to contain Pro-Gln sequences in addition to Pro-Gly sequences (Chung-Wong and Bennick, 1980). Therefore, study of the above peptides is expected to shed light on both their structural

features as well as their specific interaction with prolylhydroxylase. This in turn, would bear on the question of substrate specificity of prolylhydroxylase.

4. The selection of Boc-VPGV-OH and Boc-GVPGV-OH was based on the following reasons :

a. Valine is also one of the most frequently occurring residues in nascent procollagen chains especially in the X-position (Piez, 1976; Hoffmann *et al.*, 1980).

b. Both solution and solid-state studies have clearly demonstrated the presence of a type-II β-turn in t-Boc-Val-Pro-Gly-Val-OH (Urry *et al.*, 1974a; Yagi *et al.*, 1983). X-ray studies also suggested that the N-terminal part of the peptide molecule exhibits an extended structure followed by the β-turn with a (4 --> 1) type intra-molecular H-bond. Thus, this peptide can not only serve as a standard for β-turn but also help to examine the effect of having additional residues besides the minimum sequence for β-turn found in the tripeptide Boc-PGV-OH.

c. The addition of a Gly residue on the N-terminal side of the β-turn as in the case of t-Boc-Gly-Val-Pro-Gly-Val-OH should make it possible to compare with t-Boc-Val-Pro-Gly-Val-OH. These comparative studies may be useful in delineating the effect of adding specific residues to a β-turn containing sequence.

d. Both of the above peptides contain sequences found in elastin and have been used as elastin model peptides (Urry *et al.* 1974a, b; 1983). Elastin contains Hyp and synthetic polypeptides of these sequences were shown to undergo hydroxylation by prolylhydroxylase (Sandberg, 1976; Bhatnagar *et al.*, 1978).

These studies indicate the possibility that common structural features are recognized by prolylhydroxylase in collagen and in elastin.

The main objective at this stage was to study the conformations of the above Pro-containing oligopeptides and to correlate them with their interactions with prolylhydroxylase. The interaction could involve hydroxylation of these peptides by prolylhydroxylase and/or inhibition of the enzyme by these peptides competitively with respect to the standard substrate. Conformational studies on these peptides were discussed in detail in Chapter 5. In the following sections, the interaction of these peptides with prolylhydroxylase will be discussed.

6.4. Hydroxylation of Pro-containing Oligopeptides

As described in Chapter 2, the hydroxylation studies were carried out in two stages. In the first stage, all the peptides listed in Table 6-2 were made to interact with prolylhydroxylase under standard hydroxylating conditions, to check whether they were undergoing hydroxylation at all. The results of this initial "screening" are presented in Table 6-3. The above data indicate that all the peptides are, in fact, hydroxylated although the extent of hydroxylation varies from peptide to peptide.

Encouraged by the above results, these peptides were further studied for the determination of their kinetic parameters, namely, K_m, V_{max}, K_{cat} i.e. V_{max}/total enzyme and K_{cat}/K_m which can describe the details of enzyme-substrate interactions. However, detailed kinetic studies on Boc-PPGNHCH$_3$ were not done due to the small amounts of the peptide available. The results of the (detailed) kinetic studies on the other peptides are presented in the following subsections.

Table 6-3: Hydroxylation Data on Pro-containing Oligopeptides

Peptide[a]	% Hydroxylation[b,*]
1. Boc-PPG-$NHCH_3$	3.8± 1.20
2. Boc-PPGP-OH	6.6± 2.20
3. Boc-PPGP-$NHCH_3$	7.8± 1.80
4. Boc-PPGPP-OH	10.0± 2.50
5. Boc-PPAP-OH	2.4± 0.80
6. Boc-PPQP-OMe	1.2± 0.35
7. Boc-VPGV-OH	14.5± 2.00
8. Boc-GVPGV-OH	33.4± 7.00
9. $(Pro\text{-}Pro\text{-}Gly)_5$	100
10. poly(Pro)	0

a. Peptide concentration was 20 mM

b. Expressed with respect to $(Pro\text{-}Pro\text{-}Gly)_5$ as the standard (100% Hydroxylation); this polytripeptide undergoes hydroxylation of about 10-15% of the total hydroxylatable Pro residues under our experimental conditions.

* Mean of 3-6 repeats ± SD

6.4.1. Kinetic Data

The Lineweaver-Burk plots for the hydroxylation of different peptides are shown in Figures 6-2 to 6-8. The substrate concentration range was between 5-40 mM. Each point in these plots represent the average value of 3-4 individual trials. Table 6-4 summarizes the results of hydroxylation data in terms of the various kinetic parameters. It is seen that the K_m values varied from 10-40 mM and the V_{max} values between 1-35 μmol/h/mg.

6.5. Inhibition Studies on Prolylhydroxylase

Having shown that Pro-containing oligopeptides are hydroxylated by prolylhydroxylase, some of them were also tested for their ability to inhibit the prolylhydroxylase reaction by competing with the standard substrate, namely, $(Pro\text{-}Pro\text{-}Gly)_5$. Although, these peptides are referred to as " inhibitors" in the present studies, they are not real inhibitors in the strict sense because these peptides undergo hydroxylation by prolylhydroxylase resulting in the products of their own; a true inhibitor + enzyme complex would not break down to give product and free enzyme. However, since the hydroxylation of the standard substrate alone is monitored, any decrease in this hydroxylation due to the presence of oligopeptide substrates can be considered as "inhibition" and the peptides causing this inhibition as the "inhibitors". Two peptides were chosen for the study of inhibition :

1. t-Boc-Pro-Pro-Gly-Pro-$NHCH_3$ and

2. t-Boc-Gly-Val-Pro-Gly-Val-OH.

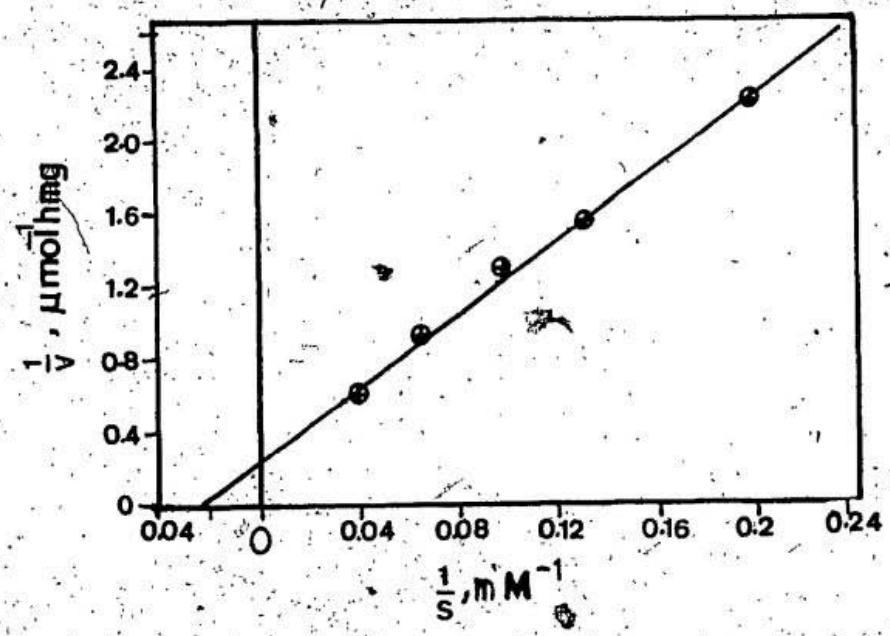


Figure 6-2: Lineweaver-Burk Plot for Hydroxylation of t-Boc-Pro-Pro-Gly-Pro-OH

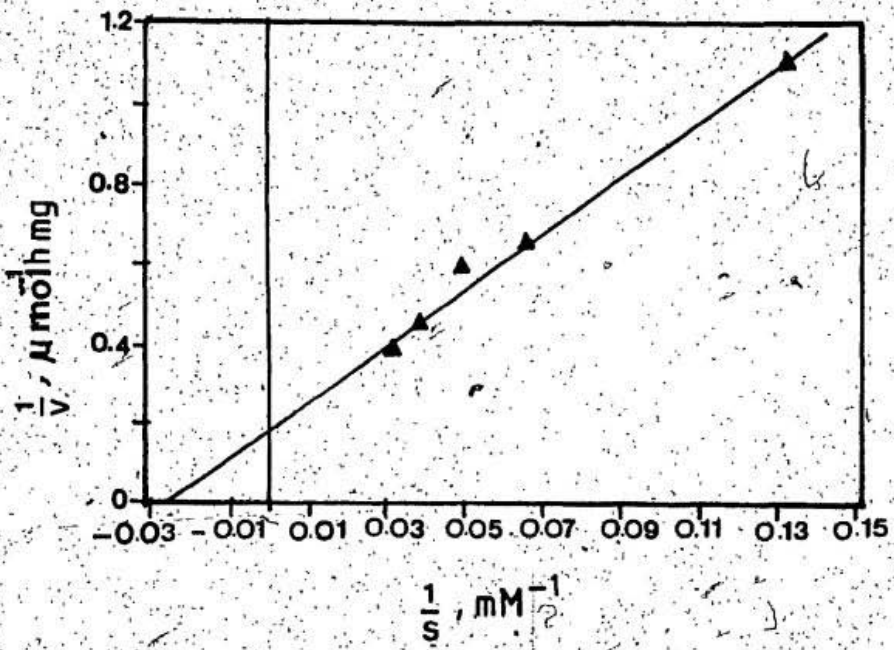


Figure 6-3: Lineweaver-Burk Plot for Hydroxylation of t-Boc-Pro-Pro-Gly-Pro-$NHCH_3$

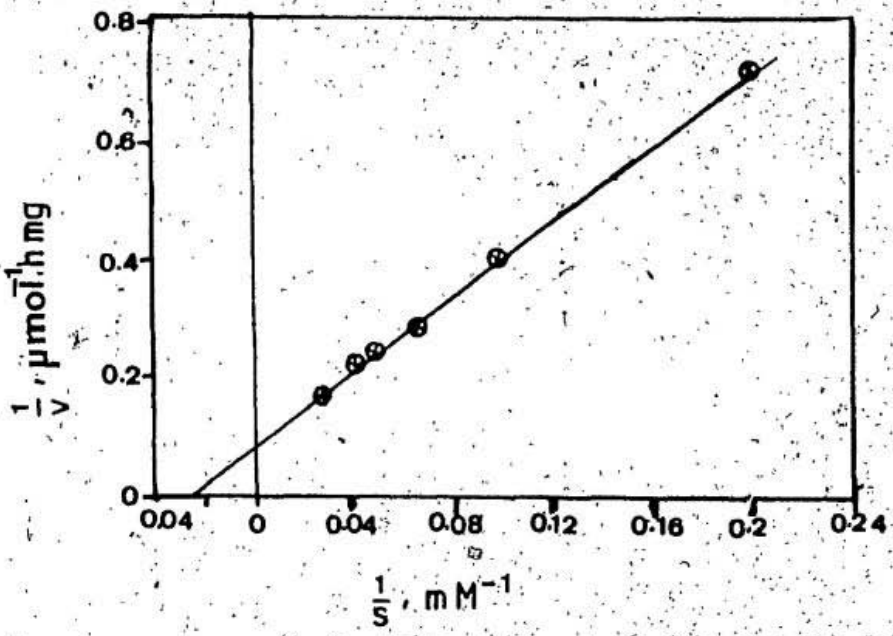


Figure 6-4: Lineweaver-Burk Plot for Hydroxylation of t-Boc-Pro-Pro-Gly-Pro-Pro-OH

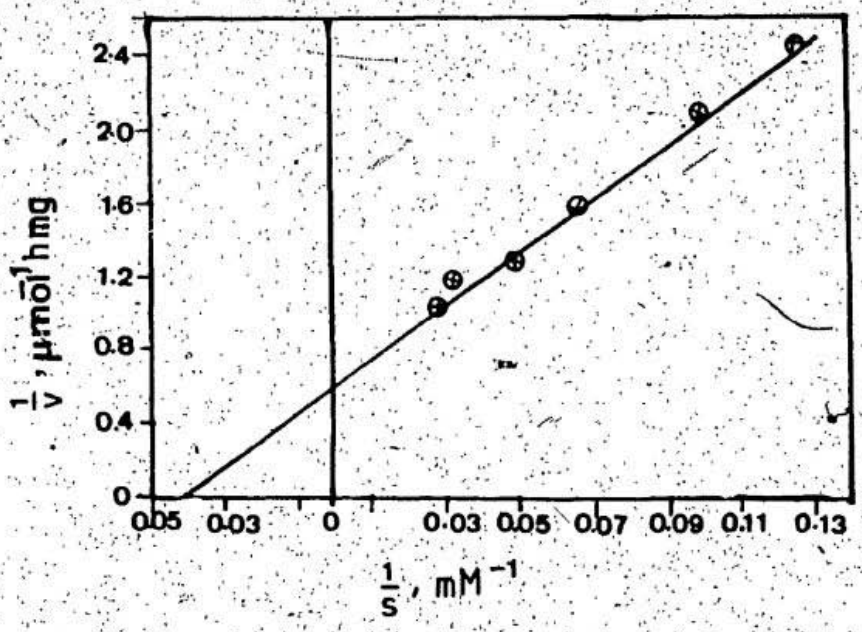


Figure 6-5: Lineweaver-Burk Plot for Hydroxylation of t-Boc-Pro-Pro-Ala-Pro-OH

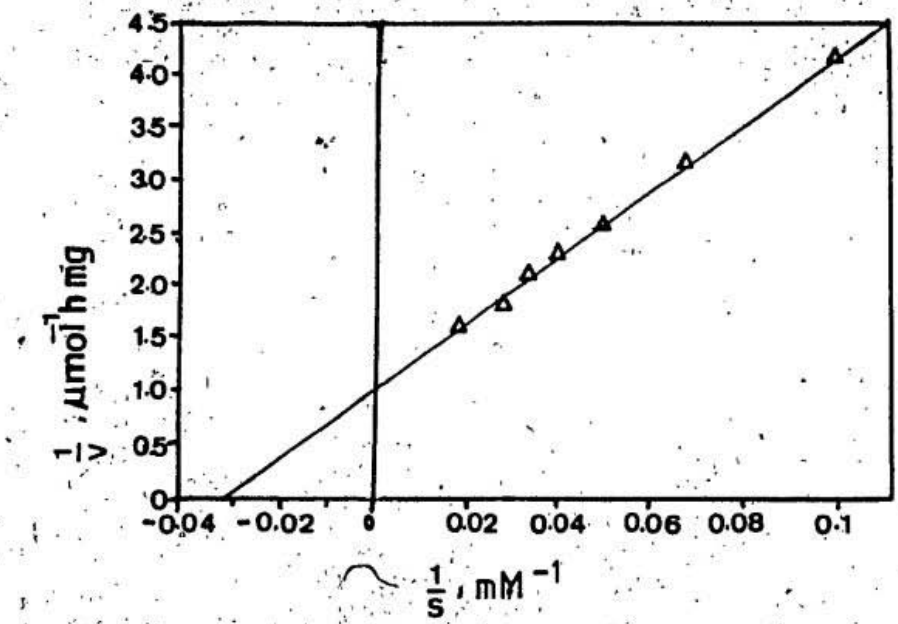


Figure 6-6: Lineweaver-Burk Plot for Hydroxylation of t-Boc-Pro-Pro-Gln-Pro-OMe

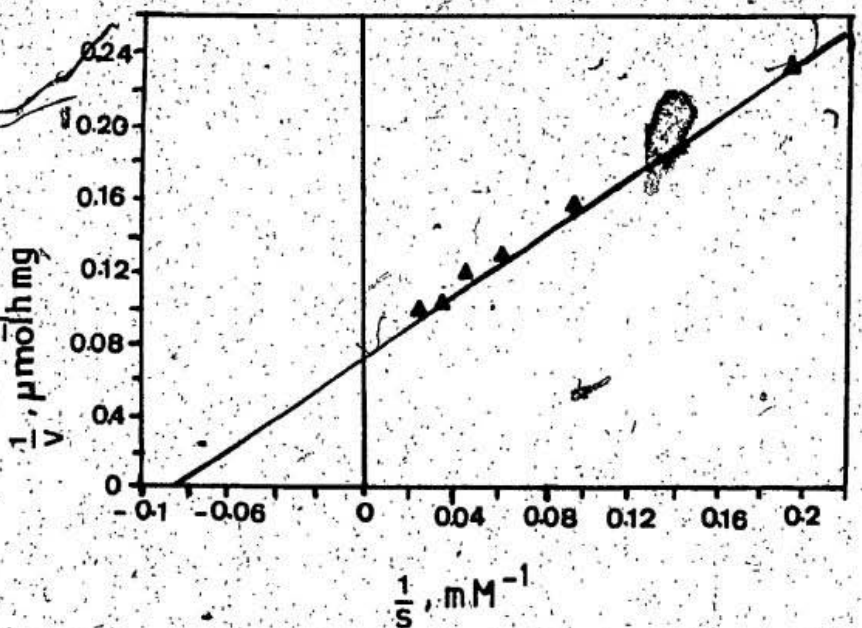


Figure 6-7: Lineweaver-Burk Plot for Hydroxylation of t-Boc-Val-Pro-Gly-Val-OH

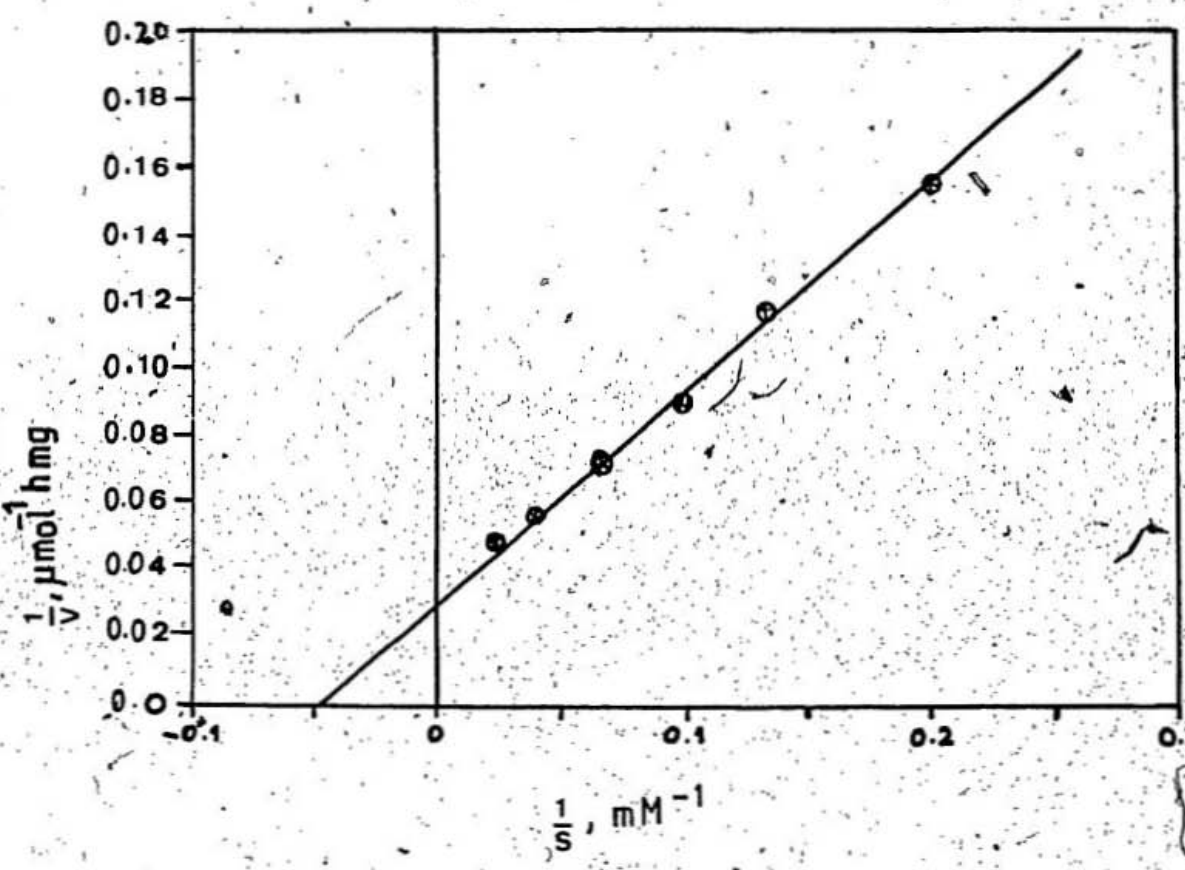


Figure 6-8: Lineweaver-Burk Plot for Hydroxylation of t-Boc-Gly-Val-Pro-Gly-Val-OH

Table 6-4: Kinetic Parameters for Hydroxylation of Pro-containing Oligopeptides

Peptides	K_m* mM	V_{max}* μmol/h/mg	K_{cat}* sec^{-1}	K_{cat}/K_m* $sec^{-1}mM^{-1}$ x 10^2
1. Boc-PPGP-OH	38.9± 0.72	4.2± 0.16	0.28±0.01	0.73± 0.02
2. Boc-PPGP-NHMe	37.2± 0.82	5.5± 0.18	0.36± 0.01	0.97± 0.03
3. Boc-PPGPP-OH	37.4± 3.42	11.9± 1.14	0.79± 0.08	2.12± 0.25
4. Boc-PPAP-OH	24.0± 0.89	1.7± 0.15	0.11± 0.01	0.46± 0.02
5. Boc-PPQP-OMe	32.2± 0.55	1.1± 0.06	0.07± 0.01	0.22± 0.01
6. Boc-VPGV-OH	10.7± 0.13	13.4± 0.23	0.88± 0.02	8.22± 0.19
7. Boc-GVPGV-OH	22.2± 1.11	35.3± 1.29	2.33± 0.01	10.5± 0.21
8. $(PPG)_5$	0.55± 0.08	110± 15	7.23± 1.10	13.3± 2.45

* Values are Mean ± S.D. ; n = 3-5 trials.

The first of these is an example of peptides that exhibited predominantly PP-II like structure in aqueous medium while the latter did not indicate this conformation explicitly in water (see Chapter 5).

As described in Chapter 2, the inhibition data were obtained using radioactively labelled $(Pro\text{-}Pro\text{-}Gly)_5$ (tritiated at position 4 of the proline rings) as the substrate to test the nature of inhibition, that is, whether competitive, uncompetitive or non-competitive. Using this method ensures that the inhibition caused by these small peptides if competitive, would be due to their binding at the same site occupied by the regular substrate on the enzyme. The inhibitor concentrations were 5 and 10 mM in the case of Boc-GVPGV-OH and 10 and 20 mM in the case of Boc-PPGP-$NHCH_3$. These ranges were chosen in view of their K_m values. The reliability of the method was checked by using the known competitive inhibitor, namely, poly(Pro) (Prockop *et al.*, 1976). Finally, the effect of the hydroxylated products on the hydroxylation of other substrates was checked using the hydroxylated counterpart of poly(Pro), namely, poly(L-Hyp) and t-Boc-Gly-Val-Hyp-Gly-Val-Pro-OH as the hydroxylated analogue of t-Boc-Gly-Val-Pro-Gly-Val-OH. This is important to know, since the inhibitor peptides are substrates by themselves and in case the products are inhibitory, then one should be able to assess the relative contribution by both the inhibitors and the products to the net inhibitory effect observed.

The results of these studies are presented in the following subsections.

6.5.1. Inhibition by Poly(Pro)

The known competitive inhibitor of prolylhydroxylase, namely, poly(Pro) was used to serve as a reference for the inhibition of $(Pro\text{-}Pro\text{-}Gly)_5$ hydroxylation by prolyhydroxylase. Poly(Pro) of an approximate M_r of 6,000 was used at a concentration of 4 μM. The concentration of $^3H(Pro\text{-}Pro\text{-}Gly)_5$ was between 0.14-1.2 mM. Figure 6-9 shows the Lineweaver-Burk double reciprocal plot for $^3H(Pro\text{-}Pro\text{-}Gly)_5$ hydroxylation data in the presence and absence of poly(Pro). The K_m value for the tritiated substrate was found to be 0.5 mM and the V_{max} value was 110 μmol/h/mg at saturating concentration of the substrate. However, in the presence of poly(Pro), while the V_{max} remained unchanged, the K_m value increased to 2.0 mM. Hence the slope of the line (K_m/V_{max}) increased from 4.5 x 10^{-3} to 18 x 10^{-3} hmg/ml. The increase would be by a factor of $[\ 1 + (I)/K_i\ ]$ in the presence of a competitive inhibitor (Cornish-Bowden, 1981). The K_i value calculated from this equation is about 1.33 μM (about 8 μg/ml) which compares well with the value reported earlier (6 μg/ml) for the poly(Pro) of similar molecular weight (Kivirikko and Prockop, 1967b, c).

6.5.2. Inhibition by Boc-PPGP-$NHCH_3$

Figure 6-10 shows the Lineweaver-Burk plot for $^3H(Pro\text{-}Pro\text{-}Gly)_5$ hydroxylation in the absence and presence of 10 and 20 mM concentration of Boc-PPGP-$NHCH_3$. In the absence of the peptide, the K_m value for $^3H(Pro\text{-}Pro\text{-}Gly)_5$ was 0.5 mM and the V_{max} value was about 110 μmol/h/mg. The slope of this line was found to be 4.4 x 10^{-3} hmg/ml, as observed earlier. However, in the presence of 10 and 20 mM of the test peptide, while the V_{max}

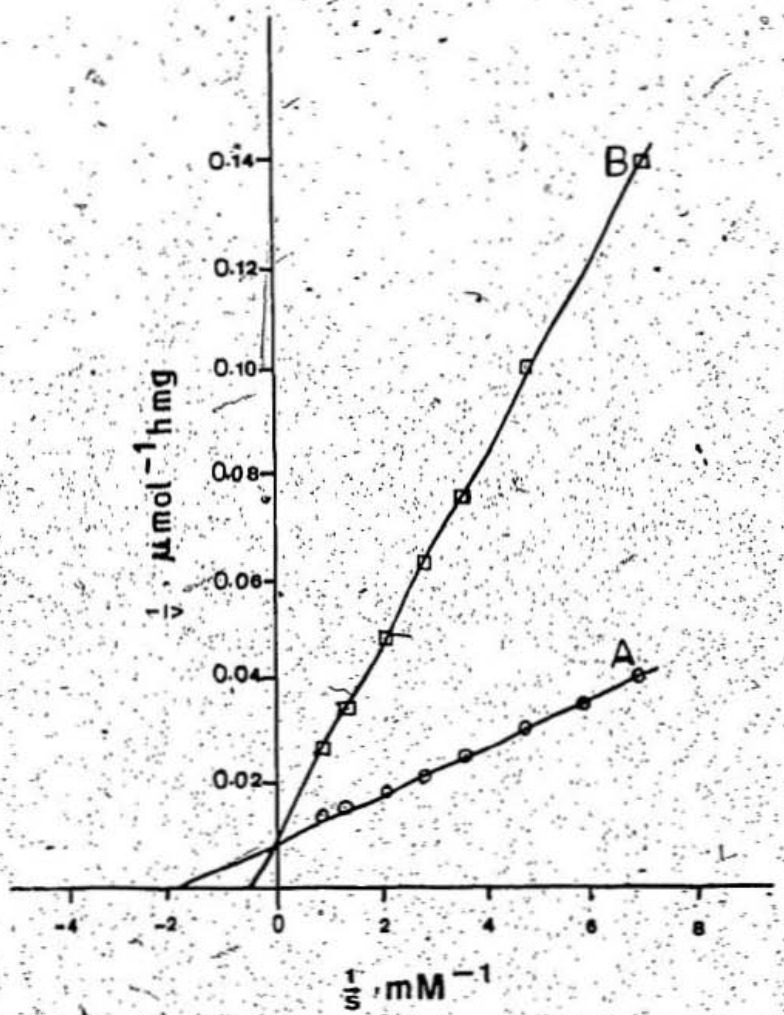


Figure 6-9: Inhibition of Prolylhydroxylase by Poly(Pro)

A = No inhibitor; B = in the presence of 4 μM poly(Pro). Each point represents the average value of atleast 3 individual trials.

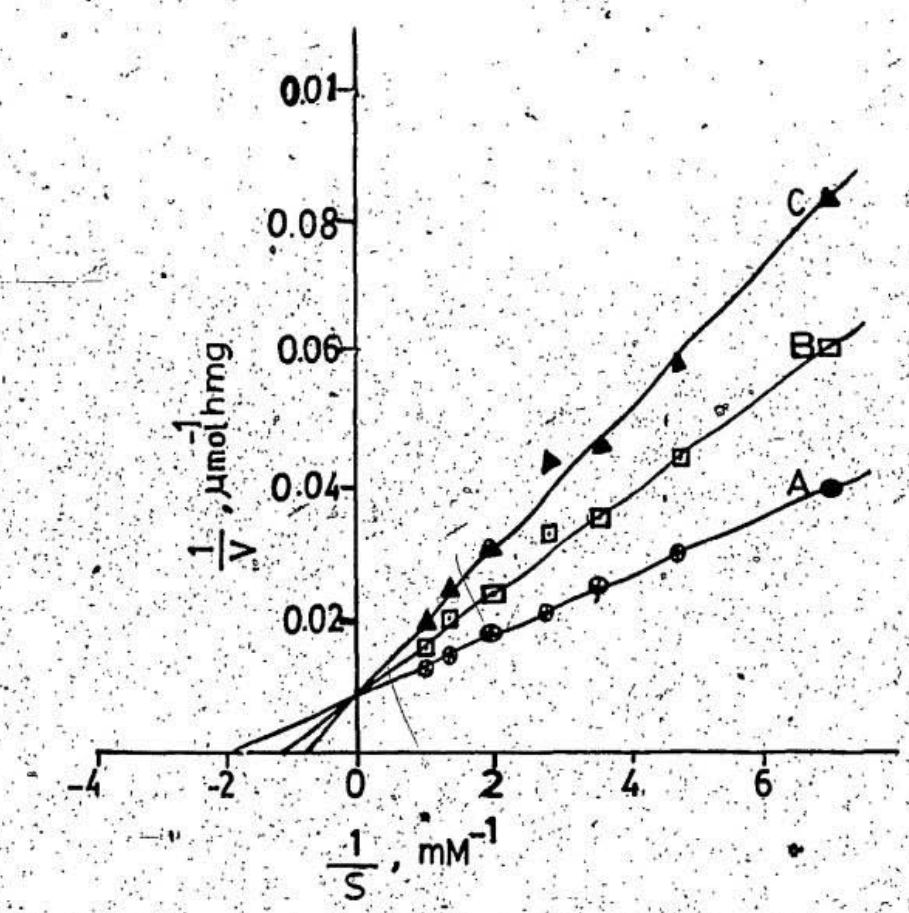


Figure 6-10: Lineweaver-Burk Plot for the Inhibition of Prolylhydroxylase by Boc-PPGP-$NHCH_3$

A = No inhibitor; B = in the presence of 10 mM inhibitor

C = in the presence of 20 mM inhibitor. Each point represents the average value of atleast 4 individual trials.

remained constant, the K_m values increased from 0.5 mM to 0.83 mM and 1.25 mM respectively. The slopes (K_m/V_{max}) also increased to 7.5 x 10^{-3} and 11.3 x 10^{-3} hmg/ml, respectively. Figure 6-11 shows the secondary plot of the slopes of the above lines *versus* the inhibitor concentration. The straight line thus obtained passed through the negative abscissa intercepting at the point equal to the K_i value (12.3 mM) for Boc-PPGP-$NHCH_3$.

Figure 6-12 shows the Dixon plot for the determination of K_i for Boc-PPGP-$NHCH_3$, where 1/v was plotted against the concentration of the inhibitor. A series of lines one at each concentration of the inhibitor were obtained that, within experimental errors, were found to intersect above the negative abscissa around 15 mM which is equal to the K_i value. This is very close to that obtained from the secondary plot described in Figure 6-11.

According to Cornish-Bowden (1981), the competitive inhibition can be confirmed by plotting s/v *versus* inhibitor concentration and obtaining a series of parallel lines. The inhibition data of Boc-PPGP-$NHCH_3$, when plotted this way, generated a series of parallel lines confirming the competitive nature of this inhibition. Figure 6-13 shows the Cornish-Bowden plot of the inhibition data for Boc-PPGP-$NHCH_3$.

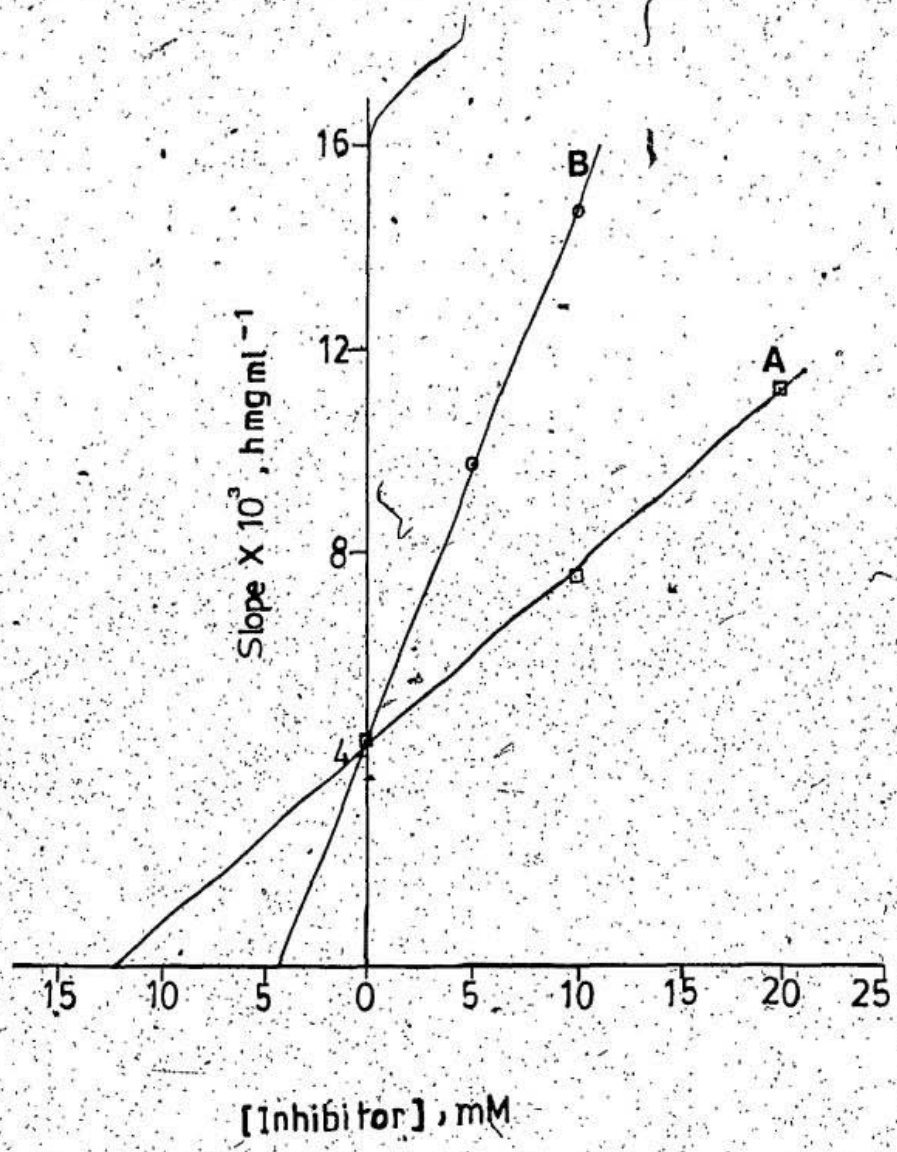


Figure 6-11: Secondary Plot for the Determination of K_i for Boc-PPGP-$NHCH_3$ and Boc-GVPGV-OH

A = Boc-PPGPNHCH$_3$; B = Boc-GVPGV-OH

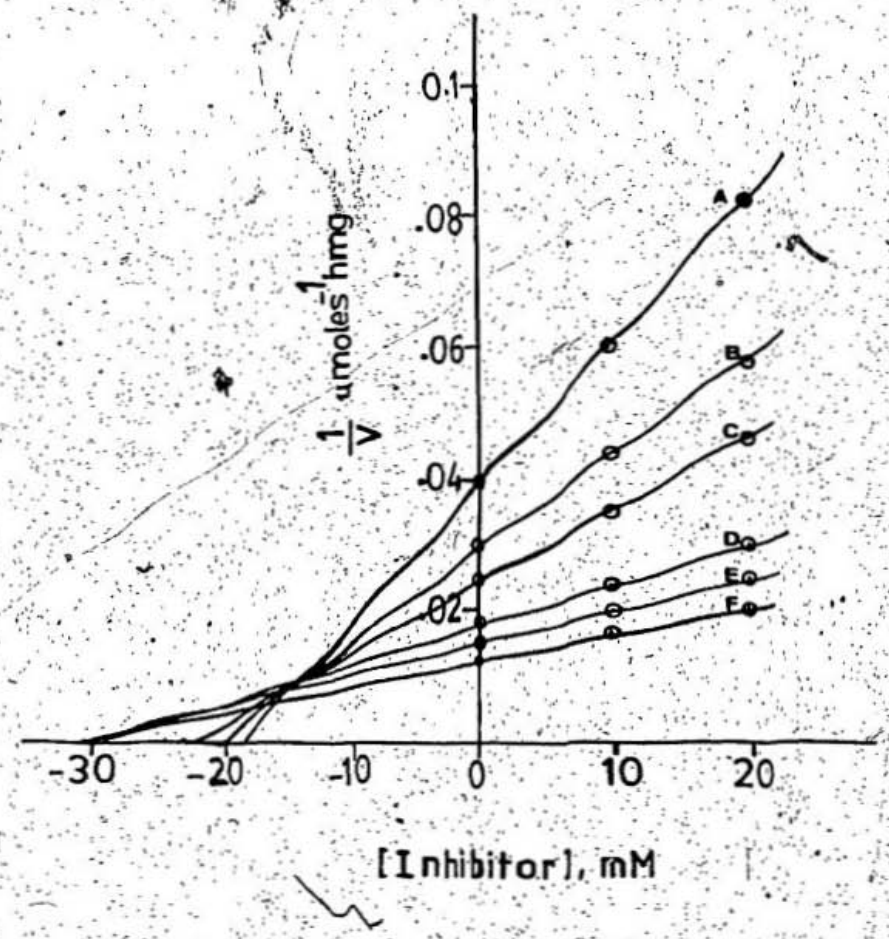


Figure 6-12: Dixon Plot for the Determination of K_i of Boc-PPGP-$NHCH_3$

at the substrate concentrations of A = 0.14 mM; B = 0.21 mM; C = 0.28 mM; D = 0.49 mM; E = 0.70 mM and F = 1.05 mM. Each point represents the average value of atleast 4 individual trials.

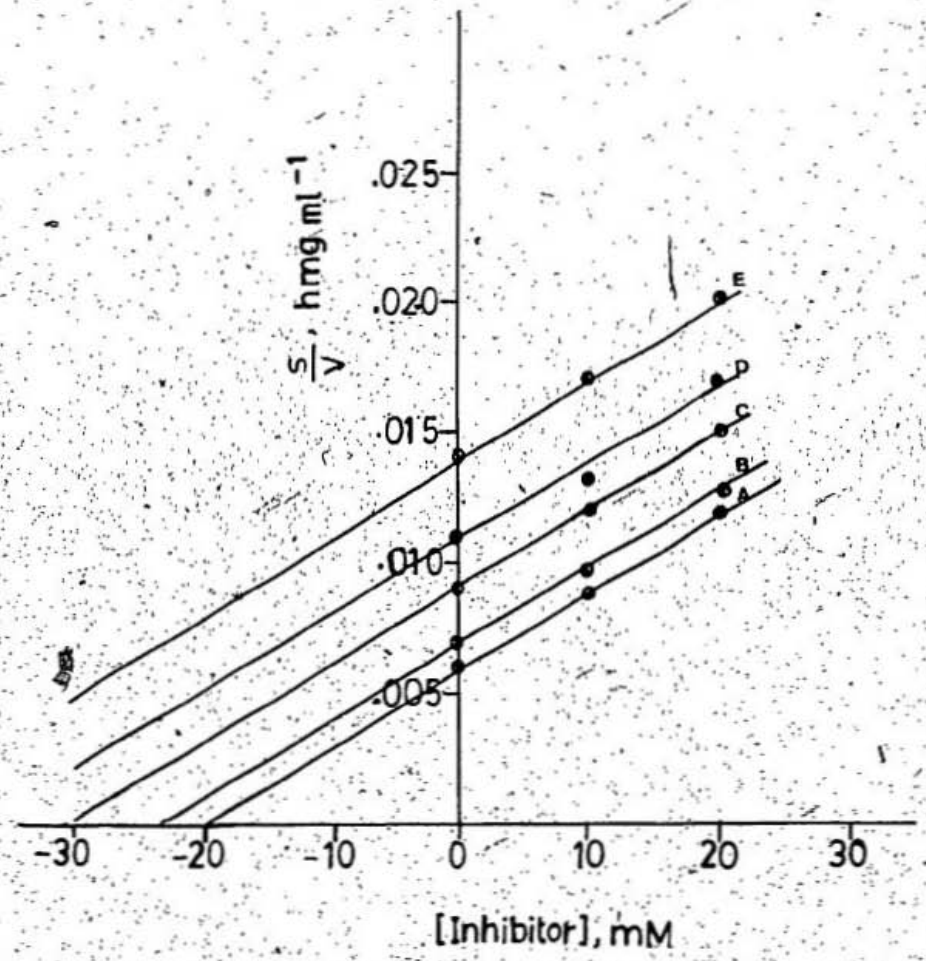


Figure 6-13: Cornish-Bowden Plot for Competitive Inhibition by Boc-PPGP-$NHCH_3$

at the substrate concentrations of A = 0.21 mM; B = 0.28 mM; C = 0.49 mM; D = 0.7 mM and E = 1.05 mM. Each point represents the average value of atleast 4 individual trials.

6.5.3. Inhibition by Boc-GVPGV-OH

Figure 6-14 shows the effect of 5 and 10 mM Boc-GVPGV-OH on the hydroxylation of $^3H(Pro\text{-}Pro\text{-}Gly)_5$ by prolylhydroxylase as represented by the double reciprocal plots. $^3H(Pro\text{-}Pro\text{-}Gly)_5$ was used in the concentration range of 0.1-1.0 mM. In the presence of 5 mM Boc-GVPGV-OH, the K_m value for $^3H(Pro\text{-}Pro\text{-}Gly)_5$ increased from 0.5 mM to 1.1 mM but the V_{max} value remained the same at 110 μmol/h/mg. The slope of the line increased from 4.4×10^{-3} to 9.8×10^{-3} hmg/ml. Increase in the concentration of Boc-GVPGV-OH to 10 mM further increased the K_m to 1.67 mM and the slope to 14.7×10^{-3} hmg/ml. Figure 6-11 shows the secondary plot of the slopes of the above lines against the inhibitor concentration. The straight line thus obtained passes through the negative abscissa at the intercept equal to the K_i value (4.5 mM).

Figure 6-15 shows the Dixon plot for the determination of K_i for Boc-GVPGV-OH, in which 1/v is plotted against the concentration of the inhibitor. A series of lines one at each concentration of the substrate was obtained that intersected above the negative abscissa at the value equal to the K_i. The value thus obtained (3.5 mM) is close to the value obtained from Figure 6-11. The Cornish-Bowden plot for the inhibition data of Boc-GVPGV-OH gave a series of parallel lines confirming the competitive nature of the inhibition. This is shown in Figure 6-16.

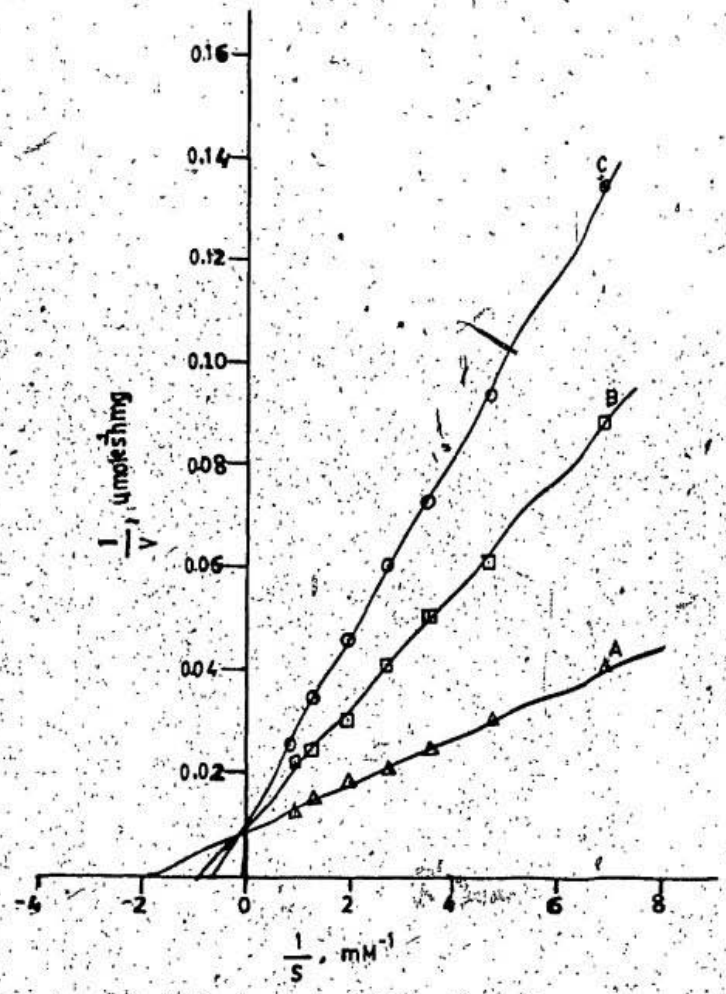


Figure 6-14: Lineweaver-Burk Plot for the Inhibition of Prolylhydroxylase by Boc-GVPGV-OH

A = No inhibitor; B = in the presence of 5 mM and C = 10 mM inhibitor. Each point represents the average of atleast 4 individual trials

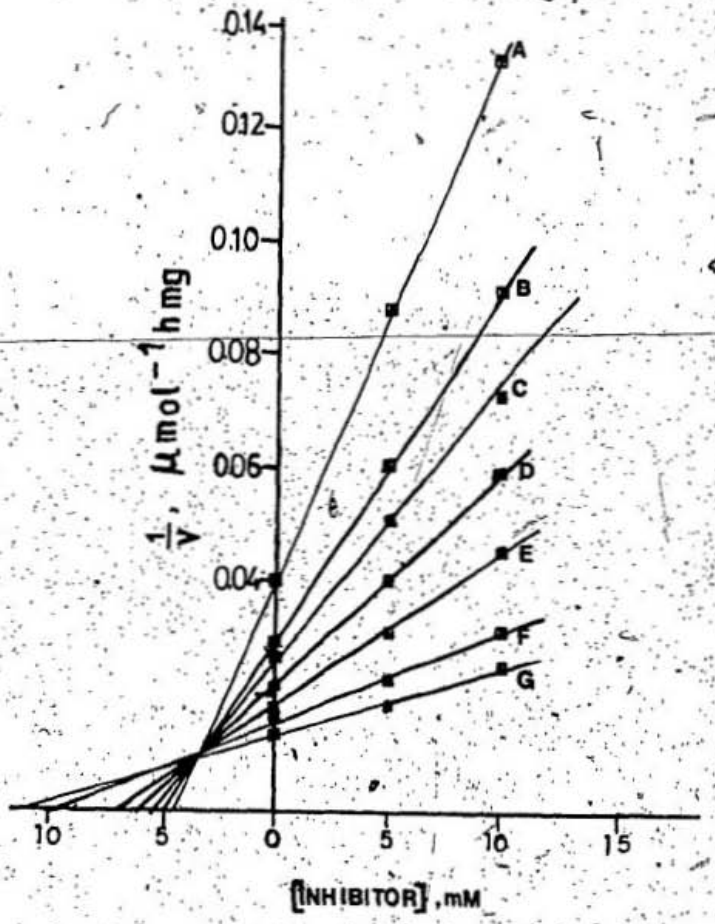


Figure 6-15: Dixon Plot for the Inhibition by Boc-GVPGV-OH at the substrate concentrations of A = 0.14 mM; B = 0.21 mM; C = 0.28 mM; D = 0.35 mM; E = 0.5 mM and F = 0.7 mM and G = 1.05 mM. Each point represents the average value of atleast 4 trials.

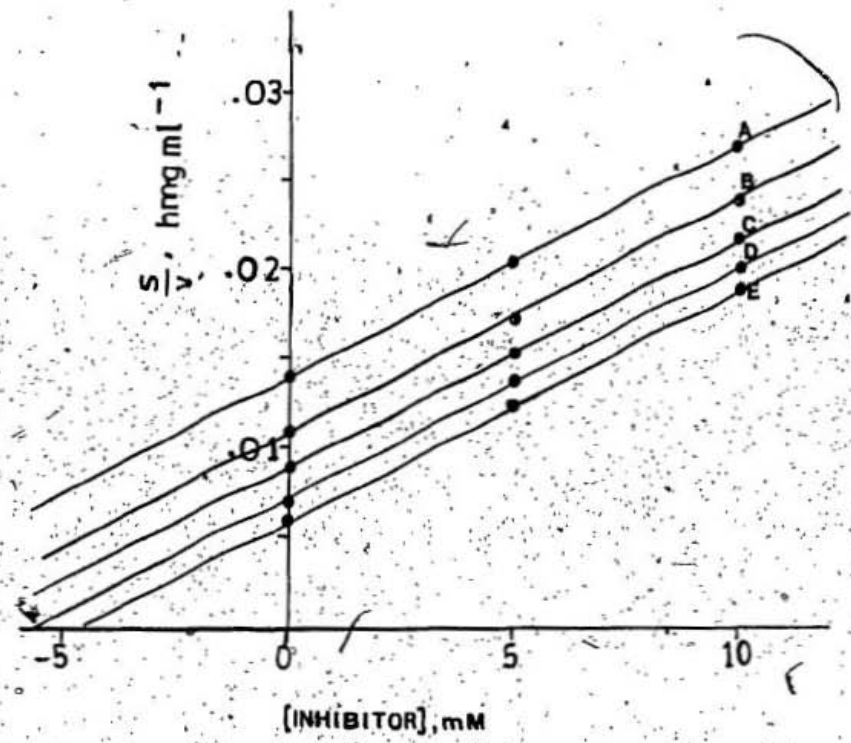


Figure 6-16: Cornish-Bowden Plot for Competitive Inhibition by Boc-GVPGV-OH

at the substrate concentrations of A = 0.14 mM; B = 0.28 mM; C = 0.49 mm; D = 0.70 mM and E = 1.05 mM. Each point represents the average value of atleast 4 trials.

6.5.4. Product Inhibition

In order to test whether the hydroxylated products exhibit any affinity towards the enzyme and bring about inhibition, product analogues, namely, Boc-GVHGVP-OH and poly(Hyp) were used. Figure 6-17 shows their effect on the hydroxylation of ^{3}H(Pro-Pro-Gly)$_5$. As seen from the figure, there is practically no inhibition either with Boc-GVHGVP-OH or with poly(L-Hyp). These results suggest that the enzyme exhibits no affinity for the product.

The results of the above inhibition studies performed with prolylhydroxylase substrates and inhibitors are summarized in Table 6-5.

6.6. Discussion

6.6.1. Hydroxylation of Peptides

The aim of the present study was to correlate the conformation of the Pro-containing peptides with their ability to act as substrates or inhibitors of prolylhydroxylase. Initial screening of the peptides (Table 6-3) indicated that all of the peptides underwent hydroxylation by prolylhydroxylase, albeit to different extents. For ease of comparison, the hydroxylation data are presented in Figure 6-18. From this figure, the relative hydroxylation of different peptides can be expressed as follows :

(Pro-Pro-Gly)$_5$ > Boc-GVPGV-OH > Boc-VPGV-OH = Boc-PPGPP-OH = Boc-PPGP-NHCH$_3$ = Boc-PPGP-OH ≥ Boc-PPG-NHCH$_3$ = Boc-PPAP-OH = Boc-PPQP-OMe

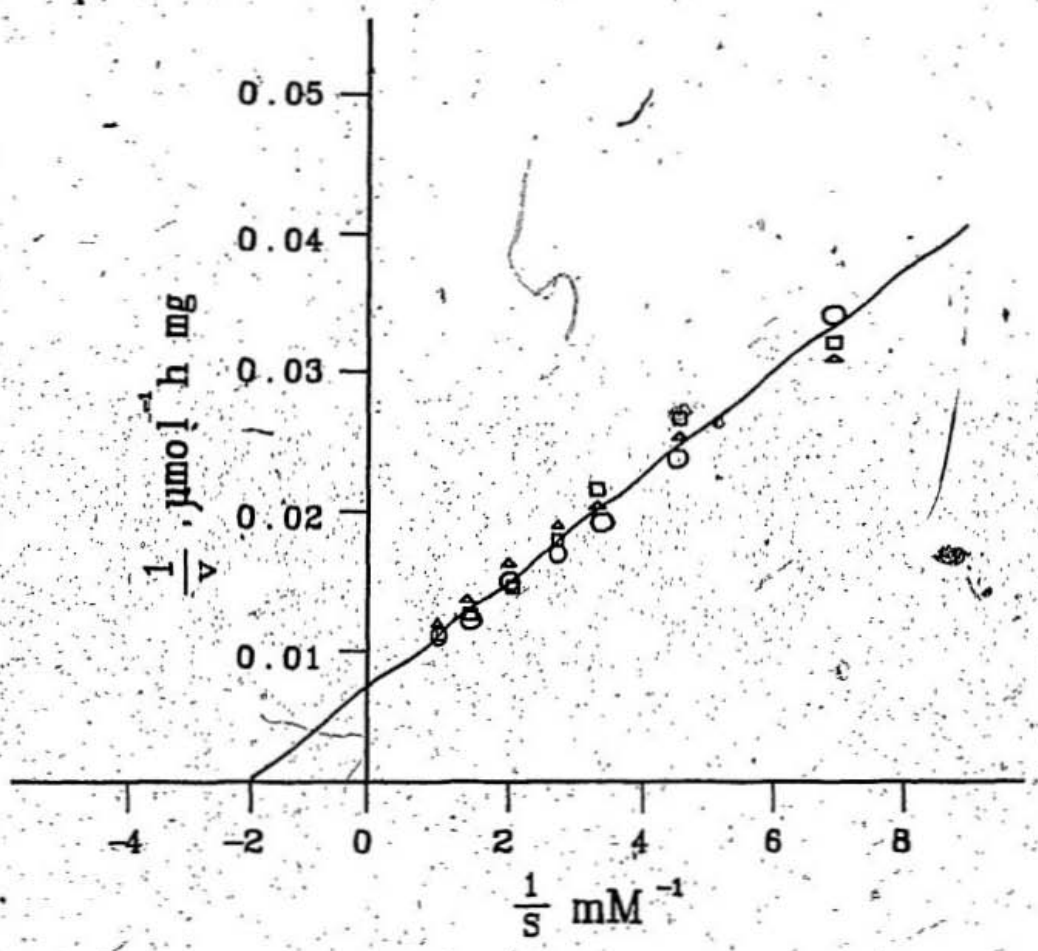


Figure 6-17: Lineweaver-Burk Plot for Product Inhibition

Triangles in the absence of the inhibitor; circles in the presence of 20 mM Boc-GVHGVP-OH and squares with 2 mg/ml poly(Hyp). Each point represents the average value of atleast 3 individual trials.

Table 6-5: Inhibition of Prolylhydroxylase by Pro- and Hyp-containing Peptides

Peptide	#Concn. mM	*K_m mM	K_m / V_{max} hmgml^{-1} x 10^3	K_i mM
1. Control $(PPG)_5$	0.14-1.05	0.5	4.4	-
2. Boc-GVPGV	5	1.10	9.8	4.3
3. Boc-GVPGV	10	1.67	14.7	4.3
4. Boc-PPGP -$NHCH_3$	10	0.83	7.5	~15
5. Boc-PPGP -$NHCH_3$	20	1.25	11.5	~15
6. Poly(Pro)II	0.004	2.00	18.2	0.0013
7. Boc-GVHyp -GVP-OH	20	0.5	4.4	-
8. Poly(Hyp)[a]	2mg/ml	0.5	4.4	-

Peptide composition was denoted in one letter code for amino acids

Inhibitor peptide concentration in the presence of varying concentrations of $(Pro\text{-}Pro\text{-}Gly)_5 \cdot 4\,H_2O$.

* K_m value of $(Pro\text{-}Pro\text{-}Gly)_5$ in the absence or presence of the inhibitor

Control $(Pro\text{-}Pro\text{-}Gly)_5$ hydroxylation in the absence of inhibitors

V_{max} = 110 μmol/h/mg in the absence and presence of the inhibitor

K_m and V_{max} values shown represent the average value of 3-5 trials

[a] Molecular weight of poly-Hyp unknown (obtained from Miles-Yeda) hence final molar concentration of the polypeptide used is unknown.

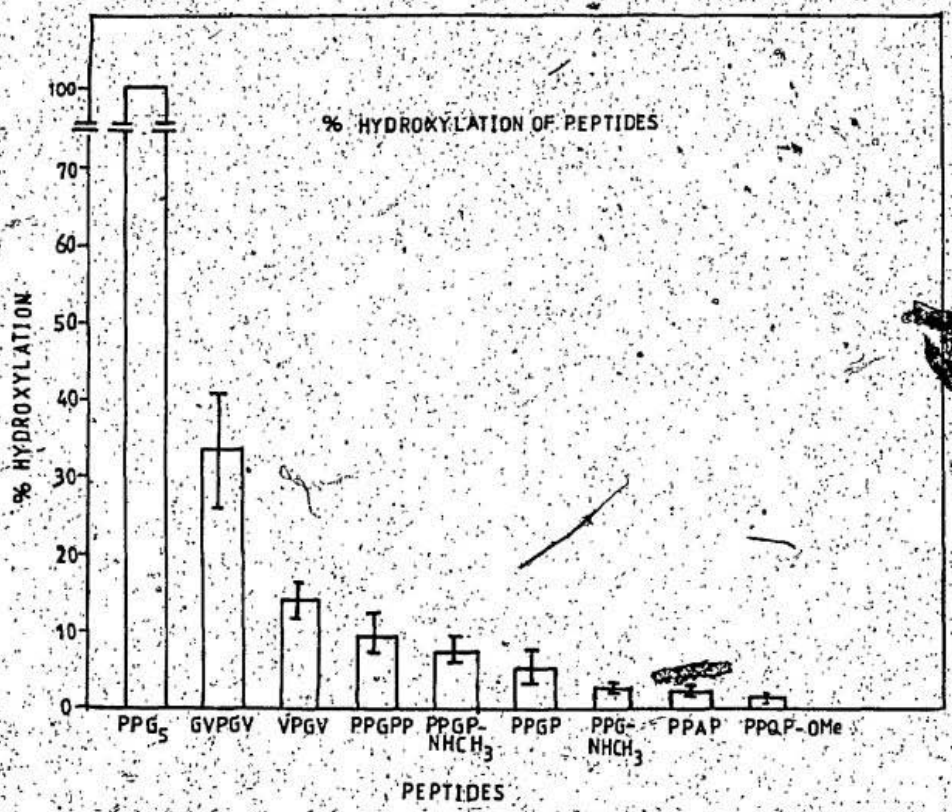


Figure 6.18: % Hydroxylation of Peptides

Although major conclusions cannot be drawn based on these data alone, it is to be noted that all the above peptides have additional residues besides the minimum tripeptide sequence needed for β-turn conformation and all of them are hydroxylated. Earlier it was observed that Boc-PGV-OH, Boc-PGA-OH and Boc-PDAA-OH (i.e. Boc-Pro-DAla-Ala-OH) can act only as inhibitors but not as substrates although they exist in β-turn conformation (Chopra and Ananthanarayanan, 1982). Together, these observations confirm that although the minimum β-turn sequence is necessary, it is not sufficient for the complete reaction; the presence and the nature of the additional residues have a role in determining the susceptibility for hydroxylation.

6.6.2. Kinetic Parameters

Although the above studies can give an idea of the relative extents of hydroxylation of the peptides, detailed studies yielding kinetic parameters are necessary in order to understand the relative affinities of binding and the relative rates with which the peptides undergo hydroxylation. This kind of information is necessary to understand which substrates can bind most efficiently and yet would undergo the catalysis less efficiently. The apparent kinetic parameters for the hydroxylation of various Pro-containing peptides are presented in Table 6-4. The apparent K_m and V_{max} values of different peptides are shown in Figures 6-19 and 6-20 for easy comparison.

As seen from the Table 6-4, $(Pro\text{-}Pro\text{-}Gly)_5$ was the 'best' substrate, with the lowest K_m value of about 0.55 mM and V_{max} of 110 μmol/h/mg which translates into a K_{cat} value of 6-8 sec^{-1}. This value is close to the value reported

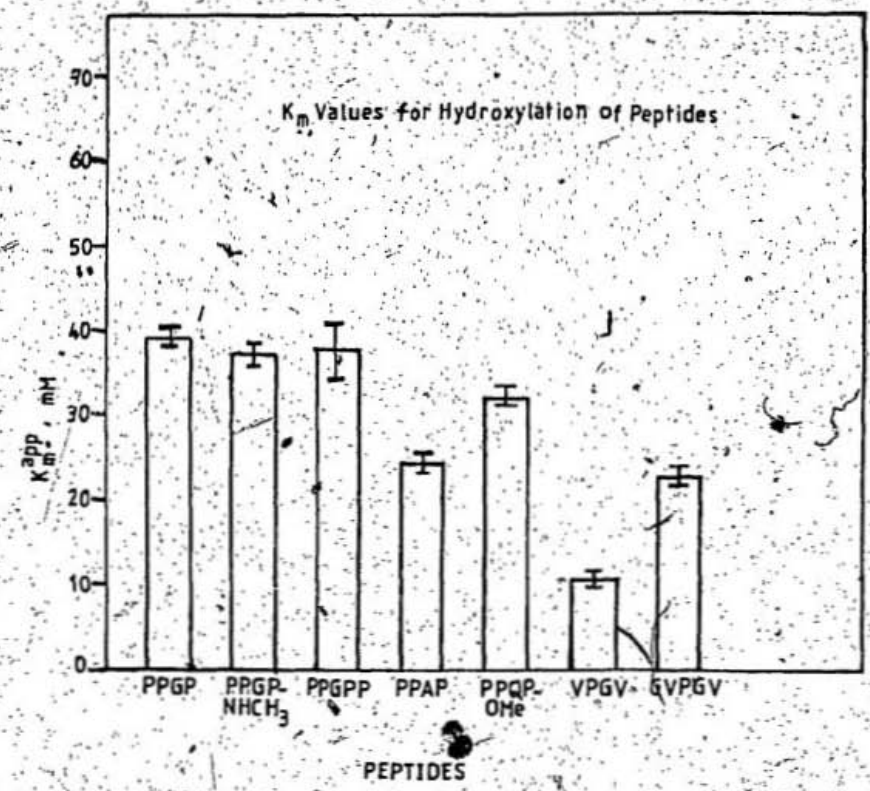


Figure 6-10: K_m Values for the Hydroxylation of Peptides

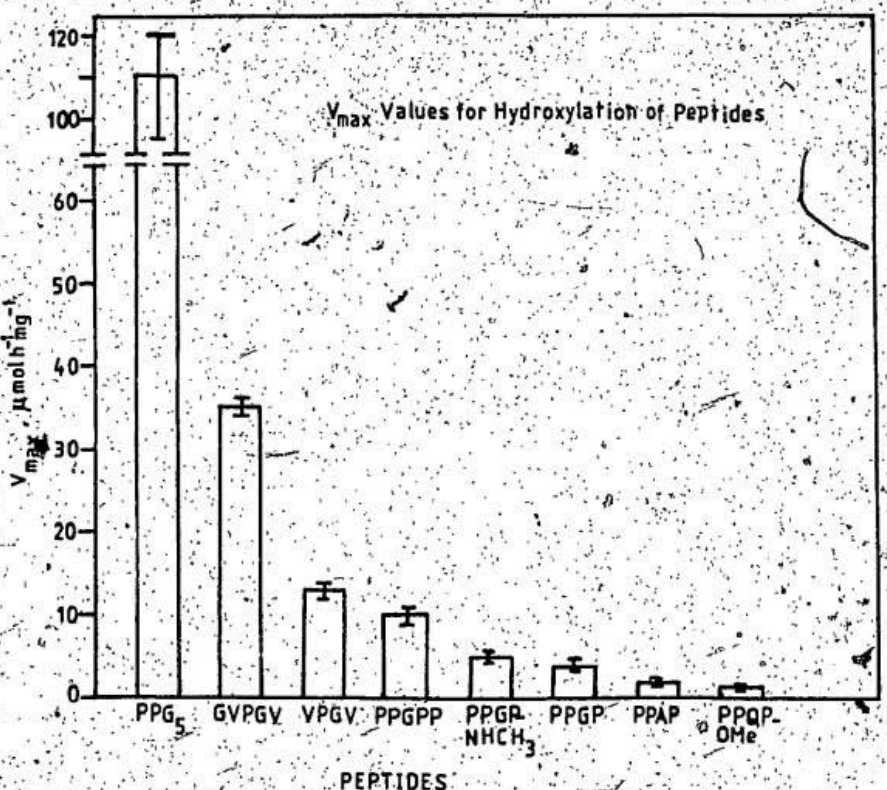


Figure 6-20: V_{max} Values for the Hydroxylation of Peptides

(4-6 sec^{-1}) by others (Prockop *et al.*, 1976). The oligopeptides, Boc-PPGP-OH and Boc-PPGP-$NHCH_3$ exhibited very similar K_m and V_{max} values. This would indicate that the conformation, at the enzyme's active site, of both Boc-PPGP-OH and Boc-PPGP-$NHCH_3$ are the same. Increasing the peptide chain length by another Pro residue on the C-terminal side as in Boc-PPGPP-OH seems to further increase the V_{max} to about 12 μmol/h/mg, although the K_m value is almost the same as above two peptides.

Substitution of Gly in Pro-Gly sequences either by Ala (as in the case of Boc-PPAP-OH) or Gln (as in the case of Boc-PPQP-OMe) seems to affect both the K_m as well as V_{max} values when compared to the above compounds. The V_{Max} values for these two compounds are very similar within the experimental error, but the K_m values seem to be significantly different. The K_m value for Boc-PPQP-OMe is about 32 mM and for Boc-PPAP-OH, it is further decreased to 24 mM, when compared to the values (about 37-39 mM) obtained for Boc-PPGP-OH, Boc-PPGP-$NHCH_3$ and Boc-PPGPP-OH.

In summary, the comparison of Boc-PPGP-OH with Boc-PPAP-OH and Boc-PPQP-OMe indicates that, although the latter two peptides exhibit better binding capabilities (as reflected by K_m values) probably due to the contacts through the sidechains, the Gly counterpart is still the better substrate in terms of V_{max}, the K_{cat} value and the ratio between the K_{cat} and K_m values.

Finally, Boc-VPGV-OH exhibits the lowest K_m (about 10 mM) among all the oligopeptides. Its K_m is about 73% lower than the proline counterpart *viz.*

extent of such a reaction. The presence and the nature of these additional residues especially on the N-terminal side of the X-Pro-Gly sequence seem to contribute to the enhanced interaction between the substrate and the enzyme. Substitution of Gly by either Ala or Gln in X-Pro-Gly sequences do not seem to abolish the interaction between the enzyme and the substrate; in fact, prolylhydroxylase seems to recognize and bind such sequences probably more effficiently than the Gly counterparts. Nevertheless, such an interaction seems to be less productive in terms of actual hydroxylation. These studies may also imply that the peptides containing Ala or Gln in the place of Gly may have common structural features that are recognized by prolylhydroxylase; this is in agreement with the structural studies presented in Chapter 5.

Based on the above results, the peptides can be arranged with respect to their degree of hydroxylation, V_{max} and K_{cat} in the order shown below :

$(Pro\text{-}Pro\text{-}Gly)_5$ > Boc-GVPGV-OH > Boc-VPGV-OH = Boc-PPGPP-OH ≥ Boc-PPGP-$NHCH_3$ = Boc-PPGP-OH > Boc-PPAP-OH = Boc-PPQP-OMe.

In contrast, there seems to be considerable variation in this order when only K_m values were considered. They are presented as follows:

Boc-PPGP-OH = Boc-PPGP-$NHCH_3$ = Boc-PPGPP-OH > Boc-PPQP-OMe > Boc-GVPGV-OH = Boc-PPAP-OH > Boc-VPGV-OH > $(Pro\text{-}Pro\text{-}Gly)_5$.

This seems to be a good example to demonstrate that considering the

apparent K_m values alone could be misleading especially when comparative analysis of homologous substrates is attempted (Fersht, 1983; Cornish-Bowden, 1981). Knowledge of other kinetic parameters is not only helpful but also necessary in delineating the relative effectiveness or susceptibility among the different substrates.

6.6.3. Inhibition of Prolylhydroxylase by Pro-containing Peptides

The inhibition studies described in the earlier section and summarized in Table 6-5 indicate that both Boc-GVPGV-OH and Boc-PPGP-$NHCH_3$ inhibit the prolylhydroxylase reaction by competing with the standard substrate $(Pro\text{-}Pro\text{-}Gly)_5$. In the presence of the inhibitory peptides, there was a rise in the K_m for $(Pro\text{-}Pro\text{-}Gly)_5$ which was dependent on the concentration of the inhibitor. However, the maximal velocity was unchanged. The slope of the line (K_m/V_{max}) also increased with increasing concentration of the inhibitor peptide suggesting that the inhibitor peptide was interacting with the same enzyme species as does the regular substrate and at the same site on the enzyme (Plesner, 1986).

Comparison of Boc-GVPGV-OH and Boc-PPGP-$NHCH_3$ indicates that the former is a better inhibitor than the latter. This is not surprising since hydroxylation data show that Boc-GVPGV-OH interacts better with the enzyme than Boc-PPGP-$NHCH_3$.

The K_i value for Boc-GVPGV-OH was found to be about 4 mM while for Boc-PPGP-$NHCH_3$ it was about three times higher. It should be noted that the K_i values obtained for Boc-GVPGV-OH and Boc-PPGP-$NHCH_3$ are much lower

than their K_m values obtained from the hydroxylation data. Ideally, one would expect that both the K_m and K_i values would be similar, if not identical, since both of them represent the equilibrium constant for the enzyme-peptide interaction, in the form of enzyme-substrate and enzyme-inhibitor complex formation, respectively. However, as pointed out by Cornish-Bowden (1981), the K_i value can be a true equilibrium-constant only if the enzyme-inhibitor complex is a dead-end complex so that it breaks down only to regenerate the enzyme and inhibitor. He also points out that in many of the more complex types of inhibition, the inhibition constant cannot be treated as a true equilibrium constant because the enzyme-inhibitor complex is not a dead-end complex. This seems to be the actual situation in the present case, where the enzyme--(Boc-GVPGV-OH) or enzyme--(Boc-PPGPNHCH$_3$) complex is not a dead-end complex but forms the products of their own. The K_m values for these reactions would be affected by the presence of the standard substrate, as would be the K_m of the standard substrate in the presence of the inhibitor peptides.

Inhibition studies by product analogues indicated that the products do not exhibit any affinity towards the enzyme; this is evident by their inability to inhibit the hydroxylation of the standard substrate. Understanding the relative affinities of the substrate and the product towards the enzyme is important since any significant affinity of the product towards the enzyme can delay or stop the release of the product from the enzyme's active site. This, in turn, would influence the turnover of the substrate molecules into the products and hence can act as a rate-determining step. Lack of inhibition, as observed in the present case, indicates the absence of affinity and hence interaction between the product and

the enzyme. Thus, the release of the product does not seem to be the rate-determining step in prolylhydroxylase reaction. This information is useful in the consideration of the mechanism of the enzymatic reaction.

6.7. Conformational Criteria for Prolylhydroxylase Reaction

As will be discussed in the present section, an insight into the exact conformational requirement for proline hydroxylation was obtained by a careful analysis of the conformation of the inhibitors and the substrates. The tripeptides t-Boc-Pro-Gly-Ala-OH, t-Boc-Pro-Gly-Val-OH and t-Boc-Pro-DAla-Ala-OH shown in Table 6.1 have the β-turn conformation (Chopra and Ananthanarayanan, 1982). In contrast, the dipeptides shown in the same Table, are found to have either a "bent" or an "open or extended" structure. Figures 6-21 and 6-22 show these structures. X-ray studies on t-Boc-Pro-Gly-OH (Benedetti *et al.*, 1977) and t-Boc-Pro-DAla-OH (Ananthanarayanan and Cameron, in press) show that the conformational angles for the Pro residues are close to those observed for the 2nd residue in type I or type II β-turn ($\phi = -60^{\circ}$ and $\psi = 150\text{-}175^{\circ}$), while those for the Gly or DAla residues are such that they would cause a "bend" similar to the one made by the 3rd residue in type I or type II β-turn, respectively ($\phi = -90^{\circ}$ or $+80^{\circ}$ and $\psi = 0^{\circ}$) (see Smith and Pease, 1980). In other words, these peptides have a 'partial β-turn' structure. In contrast, the X-ray crystal structure data on t-Boc-Gly-Pro-OH (Tanaka *et al.*, 1977) reveal an extended structure. In particular, the ϕ and ψ values for the Gly residue are high i.e. $\sim 180^{\circ}$ and $\phi_{Pro} = \sim -70^{\circ}$ and $\psi_{Pro} = +80^{\circ}$ in this compound. The conformational features of this peptide are found to be similar to that of the extended structure of PP-II. CD and IR studies on t-Boc-Gly-Pro-Pro-OH (present thesis, data not shown) also reveal a PP-II type extended structure in this peptide.

6.7.1. Correlation of the Hydroxylation Data on Pro-rich Peptides with their Conformation

As mentioned earlier, the CD data on the Pro-rich synthetic peptide substrates were interpretable in terms of the presence of different amounts of the PP-II and β-turn conformations admixed with the random-coil. Therefore, it would be of interest to correlate the hydroxylation data on these peptides with their conformational data (presented in Table 6-6).

To begin with, the "standard" polypeptide substrate $(Pro\text{-}Pro\text{-}Gly)_5$ exhibits the PP-II + β-turn structure as found from the CD data analysis and the conformational energy calculations of Lee *et al.* (1984a). The proportions of the PP-II and β-turn structures found in this peptide, *viz.* 25% and 35% respectively, might therefore be regarded as optimal for proline hydroxylation. It may be worth noting that this polypeptide also contains a considerable amount (40%) of the random coil structure. It is therefore, not made up of a regular, periodic (and hence rigid) structure but, rather, a relatively flexible one. This flexibility may indeed be necessary for the peptide substrate to be accommodated at the enzyme's active site so that minor conformational changes are still feasible.

Turning now to the oligopeptides, Boc-PPGP-OH, Boc-PPGPNHCH$_3$, Boc-PPAP-OH and Boc-PPQP-OMe; and Boc-PPGPP-OH are found to be hydroxylated to relatively smaller extent (1-10 % relative to the $(Pro\text{-}Pro\text{-}Gly)_5$). These peptides are found to contain a relatively larger PP-II content when compared to $(Pro\text{-}Pro\text{-}Gly)_5$ (see Table 6-6). This may indicate that these peptides are relatively too "rigid" when compared to the optimal conformation

Table 5-6: Correlation between Conformation and Hydroxylation of Peptides

Peptides[a]	Conformn.* f_β: f_{PPII}: f_{rc}	K_m mM	V_{max} μmol/h/mg	K_{cat} sec^{-1}	K_{cat}/K_m $sec^{-1}mM^{-1}$ x 10^2
1. Boc-PPGP-OH	35:45:20	38.9± 0.72	4.2± 0.16	0.28± 0.01	0.73± 0.02
2. Boc-PPGP-NHMe	40:40:20	37.2± 0.82	5.5± 0.18	0.36± 0.01	0.97± 0.03
3. Boc-PPGPP-OH	10:85:5	37.4± 3.42	11.9± 1.14	0.79± 0.08	2.12± 0.25
4. Boc-PPAP-OH	20:60:20	24.0± 0.89	1.7± 0.15	0.11± 0.01	0.46± 0.02
5. Boc-PPQP-OMe	20:60:20	32.2± 0.55	1.1± 0.06	0.07± 0.01	0.22± 0.01
6. Boc-VPGV-OH	65:0:35	10.7± 0.13	13.4± 0.23	0.88± 0.02	8.22± 0.19
7. Boc-GVPGV-OH	60:0:40	22.2± 1.11	35.3± 1.29	2.33± 0.01	10.5± 0.21
8. $(PPG)_5$	35:25:40	0.55± 0.08	110± 15	7.23± 1.10	13.3± 2.45

a One letter code for amino acids is used; NHMe = $NHCH_3$; Ome = OCH_3

* conformation in aqueous medium as obtained by CD spectral analysis (see Table 5-6); f_x refers to the fractional content of structure x

% Hydroxyaltion = relative to that of $(PPG)_5$

However, in view of the poor fit between the experimental and computed spectra in the case of the above oligopeptides in aqueous medium, the proportions of the β-turn and PP-II structures obtained from this analysis should be considered with caution. The improvement of the analysis (by computer) may give better values for the different secondary structures considered and hence may improve the above correlation.

6.7.2. Correlation of the Hydroxylation Data on Boc-VPGV-OH and Boc-GVPGV-OH with their Conformation

Earlier studies by Bhatnagar *et al.* (1978) showed that Pro residues in elastin-model peptide sequences, namely, $(Val\text{-}Pro\text{-}Gly\text{-}Val\text{-}Gly)_n$, $(Val\text{-}Ala\text{-}Pro\text{-}Gly\text{-}Val)_n$ and $(Val\text{-}Pro\text{-}Gly\text{-}Gly)_n$ were not only hydroxylated by prolylhydroxylase, but also inhibited the hydroxylation of the natural substrate analogue, protocollagen, in a competitive manner. These studies, therefore, strongly suggest that both (Val-Pro-Gly-Val) sequences and the substrates of prolylhydroxylase share certain common structural features. However, as discussed earlier in Chapter 5 (section 5.6.6), the IR and CD spectral characteristics of these peptides indicate an extended structure similar to that of the β-strand followed by a β-turn. The CD analysis (presented in Appendix A) also indicate the absence of PP-II structure in these peptides. These observations agree with the crystal structure data on Boc-VPGV-OH (Yagi *et al.*, 1983), where the "extended" part of the structure is seen to be made up of a β-strand like conformation, rather than PP-II like extended structure. This extended conformation is followed by a type II β-turn. Coming to the hydroxylation data, these two peptides were hydroxylated significantly by prolylhydroxylase. Thus, in

the case of these two oligopeptides, the conformational requirement for hydroxylation appeared to be an extended structure, probably β-strand, followed by a β-turn.

6.8. Conformational Model for the Substrates of Prolylhydroxylase

The above observations suggest that, in order to be a substrate for prolylhydroxylase, a peptide ought to have not only the β-turn conformation as was proposed earlier (Brahmachari and Ananthanarayanan, 1979), but in addition, an extended structure, usually the PP-II type, as well. Peptides having either one of these two required conformational features would still be "recognized" by the enzyme; however, only inhibition but not hydroxylation would result. This is obvious from the data on the β-turn tripeptides on the one hand, and the data on poly(Pro) and native collagen on the other. Interestingly, peptides that possess even a part of the β-turn or PP-II structures were also found to be inhibitory. This was seen in the case with the dipeptides Boc-Pro-DAla-OH and Boc-Pro-Gly-OH which contain the partial β-turn structures and the peptides Boc-Gly-Pro-OH which have a short stretch of the PP-II structure (see Table 5-1).

In the hydroxylated peptide substrates such as $(Pro\text{-}Pro\text{-}Gly)_5$, Boc-PPGP-OH, Boc-PPGP-$NHCH_3$ and others (except Boc-VPGV-OH and Boc-GVPGV-OH), the PP-II structure prevails at the X-Pro segment and the β-turn at the Pro-Gly segment so that the hydroxylated Pro residue would be situated at the junction between these two conformations. Therefore, from these data it may be concluded that the structural requirement for the enzymatic proline hydroxylation

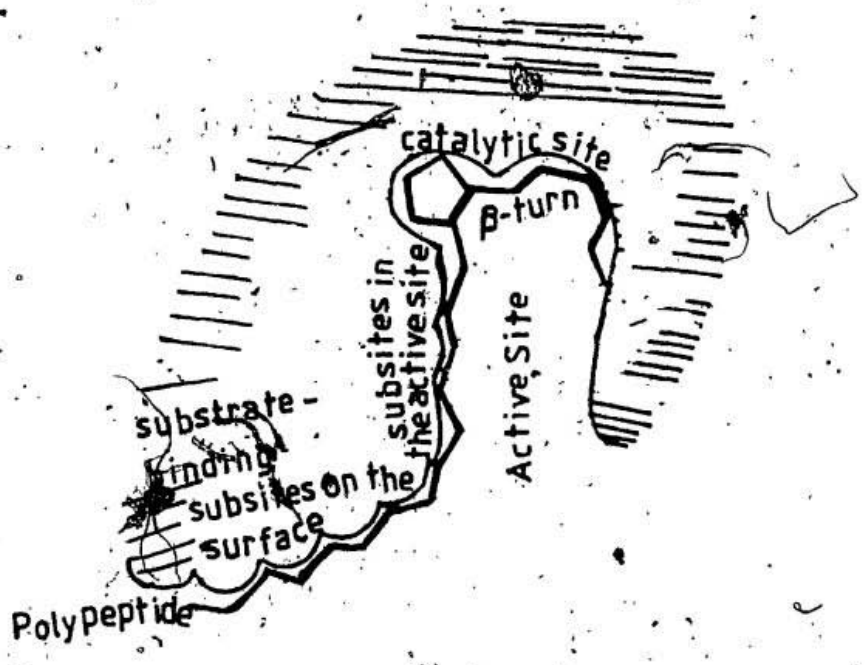
in collagen biosynthesis is the presence of the PP-II + β-turn conformation in the peptide substrate.

On the other hand, the Val-containing peptides Boc-VPGV-OH and Boc-GVPGV-OH undergo significant hydroxylation, although they do not exhibit PP-II type structure preceding the β-turn. There appears to be two possibilities to explain this observation :

1. It may be that the extended β-strand and the PP-II structures are interconvertible at the enzyme's active site i.e. the Val-Pro or Gly-Val-Pro segments of these peptides are altered from the β-strand like extended conformation to that resembling the PP-II structure. Considering that the dihedral angles for these two structures are not very different and that the energy calculations of Rapaka *et al.* (1978) indicate sufficient flexibility of the Val-Pro bond to approach the ϕ and ψ values of the Pro-Pro bond, the suggestion appears to be reasonable. Moreover, as Tiffany and Krimm (1969 a,b) demonstrated, even non-proline containing polypeptides can take up the extended, left-handed helical conformations in solution as revealed by the strong negative CD band around 200 nm. Poly(Gly) II can be cited as another example of a non-proline peptide being able to exist in an extended conformation very similar to poly(Pro) II. Interestingly, one finds the β-sheet structure as the preferred conformation for the poly(Gly) I structure (Fasman, 1967). Recent X-ray studies by Subramanian and Lalitha (1983) and the FT-IR and Raman data by Renugopalakrishnan *et al.* (1984) on a non-proline containing tripeptide Ala-Gly-Gly clearly demonstrate the conformational similarities between Ala-Gly-Gly and collagen.

2. A second possibility is that prolylhydroxylase requires some, but not a particular extended conformation followed by a β-turn in the substrate molecules. The extended conformation in the case of collagen is the PP-II structure whereas in other cases, it may be β-strand like. This can be rationalized by considering the available information on the active site geometry of prolylhydroxylase which was discussed in detail earlier in Chapter 4 (where literature references to the statements made below may be found).

In the active site of prolylhydroxylase, the substrate-binding site and the actual catalytic site are distinct but adjacent to each other. While the substrate-binding site is partly on the surface and partly at the inteior, the catalytic site is at the interior of the active site. The substrate-binding site in prolylhydroxylase tetramer (i.e. $\alpha_2\beta_2$) is present on the α-subunit whereas the β-subunit contributes to the catalytic site. There is also evidence for the presence of a number of binding subsites within the substrate-binding site. It may be that the outer substrate-binding subsites specifically bind the PP-II structure while the internal subsites near the active site may accommodate either PP-II or other extended structures such as the β-strand. The β-turn segment following this extended part in the substrate molecule would then fit into the catalytic site. The long polypeptide substrates of prolylhydroxylase bind effectively at the substrate-binding site, probably by multi-point attachment at the individual subsites. Thus, one can visualize the long nascent procollagen and its model polypeptide substrates with their PP-II type extended structure to bind effectively, at the binding subsites on the surface of the α-subunit and to extend into the catalytic site so that the β-turn structure fits into the catalytic site. This is shown in the following cartoon.



The productive binding of the polypeptide chains shown above, increases the effective concentration of the β-turn structures available for hydroxylation at the catalytic site. This would explain the usually observed lower K_m values and higher velocities in the case of longer collagen-like peptides. On the other hand, when the substrate or inhibitor is a small oligopeptide, it would not possibly bind to the entire substrate-binding region covering all the subsites and extend into the active site. In this case, the effective concentration of the substrates at the catalytic site would be governed by the distribution of these molecules over the substrate-binding site (s) on the surface and those near the catalytic site. This would be the case for the oligopeptide substrates which are predominantly in the PP-II like conformation (for example, Boc-PPGP-OH, Boc-PPGP-$NHCH_3$, Boc-PPGPP-OH, Boc-PPAP-OH and Boc-PPQP-OMe). In this situation, relatively larger amounts of the substrate are needed to increase the effective concentration

of the substrate at the active (and hence catalytic) site. This may explain the relatively higher K_m values and the lower V_{max} values observed with these peptides.

In contrast, the elastin-like oligopeptides (for example, Boc-VPGV-OH and Boc-GVPGV-OH) have a β-strand, rather than the PP-II, as the extended structure preceding the β-turn. Therefore, these molecules may be distributed more at the internal substrate-binding subsites, near the catalytic site than at the outer PP-II binding subsites which may not recognize the β-strand. The preponderence of the β-turn conformation in these peptides may also be responsible for the higher affinity they exhibit towards the catalytic site located at the interior of the molecule. Therefore, lower substrate concentrations are needed to saturate the enzyme and more substrate molecules are converted into the product molecules in a given time. This would then explain the lower K_m values and a relatively higher V_{max} values observed in the case of these elastin-like oligopeptide substrates when compared to the above Pro-rich oligopeptides.

Based on these considerations, the conformational requirements for prolylhydroxylase may be stated as follows : Prolylhydroxylase requires an extended conformation followed by a β-turn structure in its substrate molecules. In the case of procollagen, this extended structure would be that of PP-II structure. In the case of other proteins that are hydroxylated by prolylhydroxylase, other extended structures like β-strand structure may also be recognized by the enzyme, albeit to a lesser extent. The specificity of the enzyme towards the procollagen may be much higher because of the presence of PP-II like

helical structures in procollagen when compared to the other proteins, where such structures do not exist. This may be important in the selective hydroxylation of procollagen in the presence of other proteins (for example, elastin), especially, if both the substrates are present in the same cellular compartment i.e. the cisternae of endoplasmic reticulum.

Chapter 7

Conclusions

The results presented in the preceding chapters may be analyzed in the light of the objectives and scope of the present thesis, outlined in the introductory chapter. The main objectives of this thesis were four-fold:

1. Structural characteristics of, and structure-function relationship in prolylhydroxylase.

2. Conformational study of the interaction of prolylhydroxylase with its substrates, cosubstrates and cofactors.

3. Conformational characterization of synthetic Pro-containing peptides.

4. Study of the interactions of these peptides with prolylhydroxylase in biochemical as well as in conformational terms.

The conformational characterization of prolylhydroxylase and correlation of its structure with function have so far received scant attention and formed the first objective of the present thesis. My approach has been to obtain pure prolylhydroxylase from chicken embryos and characterize the enzyme in terms of secondary and tertiary structures using CD and fluorescence spectroscopy. The

same techniques were then used for studying the interaction of cofactors and cosubstrates with prolylhydroxylase. The results of these studies were presented in Chapter 4. The spectroscopic data reported here, to my knowledge, form the first report, on the conformational characteristics of purified prolylhydroxylase.

From the CD data, prolylhydroxylase seems to be rich in α-helix which comprises as much as 40% of the total secondary structural content. Also, the internal environment of the enzyme seems to have asymmetrically oriented aromatic amino acid residues. These studies are in agreement with the structural features of the β-subunit of prolylhydroxylase (Pihlajaniemi *et al.*, 1987) and protein disulphide isomerase (Edman *et al.*, 1985) as discussed in detail in Chapter 4. The effects of interactions with a peptide substrate, cosubstrate and cofactor on the conformation of prolylhydroxylase were studied. There seems to be no effect of these interactions at the secondary structure level. However, the internal environment seems to be affected due to the binding to the enzyme, of the cosubstrate (α-KG) and the cofactor ($FeSO_4$), as revealed by the concentration-dependent changes in the fluorescence spectra of prolylhydroxylase. On the other hand, interaction of substrate or substrate analogue with prolylhydroxylase did not elicit any conformational changes either at the secondary or at the tertiary structural level. Careful analysis of these observations seems to provide some important insights into the active site geometry of prolylhydroxylase and these were discussed in detail in Chapter 4. In summary, the conformational studies done on prolylhydroxylase in the present thesis seem to account for the biochemical observations made by others on the reaction mechanism of prolylhydroxylase. Moreover, these studies help to define the substrate specificity of prolylhydroxylase as discussed in Chapter 6.

The third and fourth objectives address the important question of the conformational criteria for proline hydroxylation in collagen.

At the beginning of the work reported in the present thesis, the β-turn hypothesis which implicated the β-turns as the sites of proline hydroxylation in collagen had been proposed (Brahmachari and Ananthanarayanan, 1979) and examined with a limited number of peptide models (Chopra and Ananthanarayanan, 1982). To account for certain important experimental observations (such as the inhibition of prolylhydroxylase by polyproline II and native collagen) that were not explained by the β-turn hypothesis, a new model which involves the combination of the PP-II and β-turn conformations as the requirement for proline hydroxylation was proposed by Ananthanarayanan (1984) and Ananthanarayanan *et al.*, 1985. The proposal of this model utilized some of the preliminary observations on the interaction of prolylhydroxylase with Pro-containing peptides, made by Chopra and by the present author. The stage was then set for a detailed, quantitative examination of the conformation of the tetrapeptides and pentapeptides and their interaction with prolylhydroxylase. This was expected to test the new model for its correctness and for any modifications that may be needed.

The conformation of a variety of Pro-containing peptides having different composition, sequence and length was examined in detail by spectroscopic techniques (CD and IR). The results of these studies are presented in Chapter 5. Complementing each other, the IR and CD studies clearly demonstrate the presence of an extended structure, usually the PP-II type, as well as the presence

of β-turn conformation in all of these peptides except in elastin-model peptides (Boc-VPGV-OH and Boc-GVPGV-OH). While the β-turn conformation is to be expected at the Pro-Gly sequences (Stimson *et al.*, 1977), the extended structure would be possible at X-Pro or Pro-Pro sequences (Lazarev *et al.*, 1985). The CD studies in different solvents demonstrated solvent-dependent conformational equilibria in these peptides. In non-polar solvents like TFE and dioxane, folded structures i.e. γ- or β-turns are expressed well while in polar aqueous medium, the extended PP-II like structure predominates. Increased imino acid content is found to further increase the PP-II like conformation. In elastin-model peptides, the extended conformation is found to be different from the PP-II type structure and probably similar to β-strand, followed by a β-turn (Yagi *et al.*, 1983).

Structural information on oligopeptide models of collagen has been scanty. In this context, the quantitative secondary structure analysis of Pro-containing oligopeptides using CD spectra described in the present study is useful; similar studies have not been reported earlier for small peptides, although such studies have been extensively used in the case of polypeptides and globular proteins (Greenfield and Fasman, 1969; Saxena and Wetlaufer, 1971; Chen *et al.*, 1972, 1974; Yang *et al.*, 1986). The present study indicates that this type of CD study can be useful in analyzing the proportions of different conformers present in the small peptides, at least as a first approximation. Further studies by NMR need, however, to be used to augment these observations.

The fourth objective, namely, the study of the interactions of the above model peptides with purified chicken prolylhydroxylase in biochemical as well as

conformational terms has been accomplished and in fact, forms the crux of the present thesis. These studies were aimed at providing insights into the conformational requirements of the substrate at the binding and catalytic sites of prolylhydroxylase. The results of these studies were discussed in detail in Chapter 6. According to these, the conformational requirements for the enzymatic proline hydroxylation in collagen biosynthesis is the presence of an extended conformation, usually of PP-II type, followed by a β-turn structure in its substrate molecules. Peptides with either PP-II or β-turn conformations alone acted as inhibitors but not as substrates to prolylhydroxylase since the complete structural requirement is not fulfilled in these peptide molecules. On the other hand, all the model peptides with PP-II + β-turn conformation underwent hydroxylation by prolylhydroxylase, albeit to different extents. One of the peptide substrates, namely, Boc-PPGPNHCH$_3$ with PP-II + β-turn conformation, could also competitively inhibit the hydroxylation of the standard substrate, (Pro-Pro-Gly)$_5$. The latter polypeptide substrate i.e. (Pro-Pro-Gly)$_5$ is also shown to possess PP-II + β-turn conformation (according to the analysis presented in Chapter 5 and Appendix A). These observations support the model proposed earlier by Ananthanarayanan (1984) for proline hydroxylation in collagen biosynthesis.

However, data on the two elastin-like oligopeptides Boc-VPGV-OH and Boc-GVPGV-OH offered other possibilities for the substrate specificity of prolylhydroxylase. CD studies on these peptides (Chapter 5 and Appendix A) indicated the absence of PP-II type extended structure (preceding the β-turn) in these peptides. X-ray studies on Boc-VPGV-OH indicated that the extended structure is close to that of a β-strand, followed by a β-turn (Yagi *et al.*, 1983).

These peptides undergo a significant degree of hydroxylation by prolylhydroxylase, although they have the β-strand + β-turn, rather than the PP-II + β-turn conformation in them. Both Boc-VPGV-OH and Boc-GVPGV-OH exhibited apparent K_m values much lower than the Pro-rich peptides (such as Boc-PPGP-OH, Boc-PPGPNHCH$_3$ and Boc-PPGPP-OH) and significantly higher V_{max} values. One of these peptides Boc-GVPGV-OH, also inhibited the hydroxylation of the standard substrate (Pro-Pro-Gly)$_5$, in a competitive manner. These observations could be rationalized by considering the available information on the active site geometry of prolylhydroxylase (Chapters 4 and 6).

In prolylhydroxylase, the substrate-binding site which is distinct from the catalytic site, is partly on the surface and partly at the interior, while the catalytic site is at the interior of the active site. The substrate-binding site contains a number of subsites for binding the substrate. It may be that the outer substrate-binding subsites specifically bind PP-II type structure (s), while the internal subsites near the catalytic site may accommodate either PP-II type or other extended structures such as β-strand. The β-turn segment following this extended part would then fit into the catalytic site. The long nascent procollagen and its model peptide substrates with their PP-II type extended structure bind effectively at the subsites on the surface of the enzyme and extend into the active site so that the β-turn structure fits into the catalytic site. This is shown schematically in the cartoon on page 273, 274.

On the other hand, when the substrate or inhibitor is an oligopeptide as investigated in the present thesis, the effective concentration of the substrate at

the active site would be governed by its distribution over the binding subsites on the surface and those near the catalytic site. The Pro-rich oligopeptides which are predominantly in the PP-II like structure, would bind more at the outer subsites while the elastin-like oligopeptides with a β-strand, rather than the PP-II structure, would bind more at the internal binding subsites. In this situation, relatively larger amounts of the substrate are needed in the case of Pro-rich peptides when compared to the elastin-like oligopeptides, to increase the effective concentration of the substrate at the active site. This would explain the observed K_m and V_{max} values of these oligopeptides.

Based on these considerations, the conformational requirements for prolylhydroxylase may be stated as follows : Prolylhydroxylase requires an extended conformation followed by a β-turn structure in its substrate molecules. In the case of procollagen, this extended structure would be that of PP-II structure. In the case of other proteins that are hydroxylated by prolylhydroxylase, other extended structures such as β-strand may also be recognized by the enzyme. The specificity of prolylhydroxylase towards the procollagen may be much higher because of the presence of PP-II like helical structures in procollagen, when compared to other proteins where such structures do not exist.

The above arguments in general, support the model proposed by Ananthanarayanan (1984) and in addition, offer the basis for slight modifications which would rationalize the observations made in the present thesis. Further more, they provide additional insights regarding the selective hydroxylation of

procollagen in the presence of other-proteins such as elastin, especially if both the substrates are present in the same cellular compartment (i.e. the cisternae of endoplasmic reticulum).

Backed by the data presented in this thesis, the proposed model is found to account for the available data on several peptide and polypeptide substrates or inhibitors of collagen prolylhydroxylase. Although many earlier investigators (Prockop *et al.*, 1976; Bornstein, 1974) had asserted the need for the Pro-Gly sequence in the enzymatic proline hydroxylation, an explanation in conformational terms is provided for the first time, from the work presented in this thesis. The new model provides for non-glycine residues following the Pro residue in the substrate molecule. This model also clearly points out that the specific localized conformations around the X-Pro-Gly of the substrate molecules, govern the extent of specific binding and hydroxylation of peptide substrates by prolyhydroxylase. In the hydroxylated sequences of X-Pro-Gly in substrate molecules, the effect of X residue on the extent of hydroxylation has been documented but not satisfactorily explained so far. In the light of the data presented in this thesis, it is clear that those residues in the X position whose conformational angles are conducive for the formation into an extended + β-turn conformation would favourably influence the hydroxylation while the others cannot. In their detailed stereochemical analysis of enzymatic proline hydroxylation, Hanauski-Abel and Gunzler (1982) have taken the β-turn conformation of the peptide substrate (proposed by Brahmachari and Ananthanarayanan, 1979) into consideration. It would now seem appropriate to have the PP-II + β-turn model built into such an analysis of the enzyme's

mechanism. The present studies define the conformational criteria for the enzymatic proline hydroxylation, which is a crucial event in the biosynthesis of collagen. In view of the importance of collagen in both structural integrity and functional diversity of different tissues, especially in higher vertebrates, the above studies can be considered as a positive contribution towards the understanding of biochemistry and regulation of collagen biosynthesis.

That the PP-II + β-turn structure proposed here is not unique to substrates of prolylhydroxylase alone has recently been demonstrated by Ananthanarayanan *et al.* (1987). These authors have analyzed crystal structural data on 40 globular proteins using an algorithm for detecting secondary structures and have found eight examples of a novel supersecondary structure comprised of collagen-like helix followed by β-turn. Table 7-1 shows these examples. It is important to note from the examples presented, that the presence of Pro residue does not seem to be an absolute requirement for the PP-II + β-turn structure, although five out of eight examples contain one or two Pro residues in this conformation. Earlier X-ray studies by Subramanian and Lalitha (1983) and FT-IR and Raman studies by Renugopalakrishnan *et al.* (1984) on the tripeptide Ala-Gly-Gly showed that this non-proline containing sequence, in fact, can adopt a conformation similar to that of PP-II. This highlights the fact that the presence of Pro residues may not be an absolute requirement for the formation of PP-II-like conformation. The presence of PP-II like helix followed by β-turn was also demonstrated in the avian pancreatic peptide (Blundell *et al.*, 1981) in which the PP-II like helix contains eight N-terminal residues three of which are Pro residues. This is followed by a β-turn made up of a Gly-Asp-Asp-Ala sequence.

Table 7-1: The C-βt Structure in Globular Proteins*

Protein	Residue Numbers	C-helix Sequence**	Residue Numbers	β-turn Sequence**
1. Actinidin	111-117	D-T-Y-A-N-V-P	117-120	P-Y-N-N
2. α-Chymotrypsin-A	109-115	S-T-A-A-S-P-S	115-118	S-Q-T-V
3. α-Chymotrypsin-A	119-126	S-A-V-C-L-P-S-A	125-128	S-A-A-D
4. Elastase (tosyl)	110-115	Q-S-V-T-L-N	115-118	N-S-Y-V
5. Elastase (tosyl)	119-124	V-Q-L-A-V-L-P	125-128	R-A-G-T
6. γ-IG Fab (light chain)	120-124 --	F-P-P-S-S ←	123-126 125-128	S-S-E-E E-E-L-Q
7. Parvalbumin	73-79	D-A-R-A-L-T-D	78-81	T-D-G-E
8. Rhodanase	190-195	T-Q-P-E-P-D	193-196	E-P-D-A

* Based on analysis of the crystal structure data on 40 proteins.

** Amino acids are denoted in one-letter code; C-helix = Collagen-helix;

C-βt = Collagen-helix followed by β-turn.

Adapted from Ananthanarayanan *et al.*, 1987.

7.1. General Perspective

I believe that the present studies project a clear picture of the substrate specificity of prolylhydroxylase in definitive conformational terms and the model proposed seems to account for all the available data in this regard. As suggested earlier by Brahmachari and Ananthanarayanan (1979) and Ananthanarayanan *et al.* 1985), the β-turn conformation may be involved as a general recognition site for enzymes involved in posttranslational modifications of proteins and peptides. In this context, it may be pertinent to recall the hydroxylation studies on bradykinin. McGee *et al.* (1971) found that bradykinin and its analogues were significantly hydroxylated by prolylhydroxylase. This nonapeptide has the sequence: Arg-Pro-Pro-Gly-Phe-Ser-Pro-Phe-Arg. The presence of the β-turn conformation in the Pro-Gly-Phe segment is strongly suggested by the observation of a type II β-turn in N-Ac-Pro-Gly-Phe-OH by X-ray crystallography (Brahmachari *et al.*, 1981). It is interesting to note that this segment is preceded by a Pro residue so that one could expect the PP-II + β-turn conformation in the Pro-Pro-Gly-Phe segment of bradykinin. This would then account for the observed hydroxylation of bradykinin indicating that the complete structural requirement for enzymatic proline hydroxylation is met by its sequence (rather than the presence of the β-turn at Pro-Gly-Phe segment alone). It should be interesting to see if the crystal structure of bradykinin, which is not yet available, would bear out this speculation.

Similarly, studies on a Pro-rich protein from human saliva also seem to support the above model. The complete amino acid sequence of a Pro-rich

salivary protein (Bennick, 1975) has recently been determined (Chung-Wong and Bennick, 1980) and it contains the following Pro-containing sequences :

Pro-Pro-Gly-Lys; Pro-Pro-Leu-Gly; Pro-Pro-Gln-Gly; Pro-Pro-Arg-Gly; Pro-Pro-Gln-Gln; Arg-Pro-Gly-Gly; Val-Pro-Leu-Val; His-Pro-Arg-Pro; Lys-Pro-Glu-Gly and Gln-Pro-Ser-Ala.

Among these, Pro-Gly sequences occur only twice while Pro-Gln sequence occurs about 15 times and mostly in the tetrapeptide sequences of Pro-Pro-Gln-Gln or Pro-Pro-Gln-Gly. It is possible that these Pro- containing sequences may adopt the PP-II + β-turn structure as seen in Boc-Pro-Pro-Gly-Pro-OH and Pro-Pro-Gln-Pro-OMe (Table 5.6). This protein was generously supplied by Dr. Bennick of the University of Toronto and I have found that this protein is hydroxylated by prolylhydroxylase (about 2-4% relative to $(Pro\text{-}Pro\text{-}Gly)_5$). This can be explained in the light of the model presented above if we assume that these and other Pro-rich sequences adopt the PP-II structure followed by a β-turn conformation and hence are recognized by prolylhydroxylase. In this context, it is pertinent to recall the conformational and hydroxylation data on Boc-Pro-Pro-Gln-Pro-OMe sequence presented in Chapter 6.

Studies on c-AMP-dependent protein kinases using model substrates ike casein (Kemp *et al.*, 1975), lysozyme (Bylund and Krebs, 1975), protamines (Shenolikar, 1978) and synthetic peptides (Feramisco, 1980) and on c-AMP-independent protein kinases using denatured pepsin and antiprotease C-II as the substrates (Meggio *et al.*, 1981) showed that the predicted secondary structure

repeatedly occuring at the phosphorylation sites is the β-turn which apparently includes the target residue i.e. Thr or Ser. However, these studies are not considered to be definitive since structural information was sought using predictive methods alone (See Smith and Pease, 1980; Rose *et al.*, 1985). More direct evidence for the recognition of β-turn peptides was reported by Tinker *et al.* (1986). These studies demonstrated that the phosphorylation (*in vitro*) by a protein kinase, of tyrosyl residues contained in characterized β-turn sequences. A recent report by Zijenah and Ananthanarayanan (1987) and Takahashi *et al.* (1987) also implicates the β-turn as the recognition site for the hydroxylation of Lys-containing peptide substrates by lysylhydroxylase. These studies offer strong support for the earlier speculation by Brahmachari and Ananthanarayanan (1979) that β-turn conformation could be the recognition site for many post-translational modifications including phosphorylation and glycosylation, in addition to, proline hydroxylation.

It is interesting to note in this context, that type I collagen from bovine dermis is phosphorylated by c-AMP-dependednt protein kinase from bovine heart, but only in the non-helical form obtained through thermal denaturation and not in the triple-helical conformation (Glass and McPherson, 1986). It is worth recalling that prolylhydroxylase and lysylhydroxylase do not act upon the triple-helical native collagen but further, hydroxylate Pro or Lys residues when this protein is heat-denatured (Prockop *et al.*, 1976) to the non-helical form. Therefore, the same or similar secondary structural features may be recognized by the protein kinase and by prolyl and lysyl hydroxylases.

Finally, the concluding remark is, as always, "Conformation is Information".

Chapter 8

References

1. Adler, A.J., Greenfield, N.J. and Fasman, G.D. (1973) Methods in Enzymol. 27, 675-745.
2. Ananthanarayanan, V.S. (1983a) In *Conformation in Biology, The Ferstchrift* (Eds. Srinivasan R. and Sharma, R.H.), pp 99-111, Adenine Press, New York.
3. Ananthanarayanan, V.S. (1983b) J. Biomol. Str. & Dyn. 1, 843-855.
4. Ananthanarayanan, V.S. (1984) In *Proceedings of the 16th FEBS Congress: Part A* (Eds. Ovchinnikov, Yu. A.), pp 23-30, VNU Press, The Netherlands.
5. Ananthanarayanan, V.S. and Cameron, T.S. (1988) Int. J. Pept. Protein Res. 29, 000-000 (in Press).
6. Ananthanarayanan, V.S. and Shyamasundar, N. (1981) Biopolymers 102, 295-301.
7. Ananthanarayanan, V.S., Brahmachari, S.K., Rapaka, R.S. and Bhatnagar, R.S. (1976) Biopolymers 15, 707-716.
8. Ananthanarayanan, V.S., Brahmachari, S.K. and Pattabhiraman, N. (1984) Arch. Biochem. Biophys. 232, 482-495.
9. Ananthanarayanan, V.S., Attah-Poku, S.K., Mukkamala, P.L. and Rohse, P.H. (1985) J. Biosci. 7, 001-013.
10. Ananthanarayanan, V.S., Soman, K.V. and Ramakrishnan, C. (1988) J. Mol. Biol. 198, 705-710.
11. Anglister, L., Rogozinski, S. and Silman, I. (1976) FEBS Letters 69, 129-132.
12. Aubert, J.P., Boersma, A., Loucheux-Lefebvre, M.H., Degand, P. and Biserti, G. (1975) FEBS Letters 56, 263-267.
13. Aubry, A., Protas, J., Boussard, G. and Marraud, M. (1977) Acta Cryst. Sect. B33, 2399-2406.
14. Aubry, A., Protas, J., Boussard, G. and Marraud, M. (1979) Acta Cryst. Ser B 35, 694-699.
15. Aubry, A., Vitoux, H.B. and Marraud, M. (1985) Biopolymers 24, 1089-1100.
16. Avignon, M., Huong, P.V. and Lascombe, J. (1969) Biopolymers 8, 69-89.
17. Ayato, H., Tanaka, I. and Ashida, T. (1980) J. Am. Chem. Soc. 103, 5902-5905.

18. Ayato, H., Tanaka, I. and Ashida, T. (1981) J. Am. Chem. Soc. 103, 6869-6873.
19. Bachinger, H.P., Bruckner, P., Timpl, R. and Engel, J. (1978) Eur. J. Biochem. 90, 605-613.
20. Balazs, E.A. Ed. (1970) In *Chemistry and Molecular Biology of the Intercellular Matrix*, Vols. 1-3, Academic Press, New York.
21. Bandekar, J. and Krimm, S. (1986) Adv. in Prot. Chem. 38, 269-364.
22. Bandekar, J., Evans, D.J., Krimm, S., Leach, S.J., Lee, S., Mc Quie, J.R., Minasiam, E., Nemethy, G., Pottle, P.S., Scheraga, H.A., Stimson, E.R. and Woody, R. W. (1982) Int. J. Pep. Prot. Res. 19, 187-205.
23. Bansal, M. and Ananthanarayanan, V.S. (1987) in Press and appeared in (1988) Biopolymers 27, 299-312.
24. Bansal, M., Ramakrishnan, C. and Ramachandran, G. N. (1975) Proc. Ind Acad. Sci. USA Vol. LXXXII, Section 'A', No.4, 152-164.
25. Bansal, M., Brahmachari, S.K. and Sasisekharan, V. (1979) Macromolecules 12, 19-22.
26. Bayley, P. (1980) In *An Introduction to Spectroscopy for Biochemists*, (Ed. Brown, S.B.) pp 148-224, Academic Press, New York.
27. Benedetti, E. (1977) in "Peptides", Proceedings of the Fifth American Peptide Symposium (Eds. Goodman, M. and Meienhofer, J.) pp257-271, John Wiley and Sons, New York.
28. Benedetti, E., Pavone, V., Toniolo, C. and Bonora, G.M. (1977) Macromolecules 10, 1350-1356.
29. Benedetti, E., Bovoso, A., Di Blasio, B., Pavone, C., Pedone, C., Toniolo, C. and Bonora, G.M. (1983) Biopolymers 22, 305-317.
30. Bennick, A. (1975) Biochem. J. 145, 557-567.
31. Berg, R.A. and Prockop, D.J. (1973a) J. Biol. Chem. 248, 1175-1182.
32. Berg, R.A. and Prockop, D.J. (1973b) Biochem. Biophys. Res. Commun. 52, 115-119.
33. Berg, R.A. and Prockop, D.J. (1973c) Biochemistry 12, 3395-3401.
34. Berg, R.A., Kishida, Y., Sakakibara, S. and Prockop, D.J. (1977) Biochemistry 16, 1615-1620.
35. Berg, R.A., Kedersha, N.L. and Guzman, N.A. (1979) J. Biol. Chem. 254, 3111-3118.
36. Berger, A. and Schechter, I. (1970) Philos. Trans. R. Soc. London B257, 249-264.
37. Bhatnagar, R.S. and Rapaka, R.S. (1975) Biopolymers 14, 597-603.
38. Bhatnagar, R.S. and Rapaka, R.S. (1976) In *Biochemistry of Collagen* (Eds. Ramachandran, G.N. and Reddi, A.H.), pp 479-523, Plenum Press, New York.
39. Bhatnagar, R.S., Rapaka, R.S. and Urry, D. W. (1978) FEBS Letters 95, 61-64.
40. BioRad Standard Assay (1977) Procedure supplied with the Assay Kit.

41. Blout, E.R., Deber, C.M. and Pease, L.G. (1974) In *Peptides, Polypeptides and Proteins*, (Eds. Blout, E.R., Bovey, F.A., Goodman, M. Lotan, S.) pp 266-281.

42. Blundell, T.L., Pitts, J.E., Tickle, I.J., Wood, S.P. and Wu, C.W. (1981) Proc. Natl. Acad. Sci. USA 78, 4175-4179.

43. Bodanski, M. and Bodanski, S. (1984) *Principles of Peptide Synthesis*, Springer-Verlog, New York.

44. Bornstein, P. (1967a) J. Biol. Chem. 242, 2572-2574.

45. Bornstein, P. (1967b) Biochemistry 6, 3082-3093.

46. Bornstein, P. (1974) Ann. Rev. Biochem. 43, 567-603.

47. Bornstein, P. and Traub, W. (1979) In *Proteins*, (The Chemistry and Biology of Collagen). Vol. IV. (Eds. Neurath, H and Hill, R.), p 411-632, Academic Press, New York.

48. Boussard, G, Marraud, M. and Aubry, A. (1979) Biopolymers 18, 1297-1331.

49. Brahmachari, S.K. and Ananthanarayanan, V.S. (1979) Proc. Natl. Acad. Sci. USA 76, 5119-5123.

50. Brahmachari, S.K., Ananthanarayanan, V.S., Rapaka, R.S. and Bhatnagar, R.S. (1978). Biopolymers 17, 2097-2105.

51. Brahmachari, S.K., Ananthanarayanan, V.S., Brahms, S., Brahms, J., Rapaka, R.S. and Bhatnagar, R.S. (1979) Biochem. Biophys. Res. Commun. 86, 605-612.

52. Brahmachari, S.K., Bhat, T.N., Sudhakar, V., Vijayan, M., Rapaka, R.S., Bhatnagar, R. and Ananthanarayanan, V.S. (1981) J. Am. Chem. Soc. 103, 1703-1708.

53. Brahmachari, S.K., Bhatnagar, R.S. and Ananthanarayanan, V.S. (1982) Biopolymers 21, 1107-1125.

54. Brahms, S. and Brahms, J. (1980) J. Mol. Biol. 138, 149-178.

55. Brahms, S., Brahms, J., Spach, G. and Brack, A. (1977) Proc. Natl. Acad. Sci. USA 74, 3208-3212.

56. Brandts, J.F., Halvorson, H.R and Brenner, M. (1975) Biochemistry 14, 4953-4963.

57. Brodsky-Doyle, B., Leonard, K.R. and Reid, K.B.M. (1976) Biochem. J 159, 279-286.

58. Brown, F.R., III, Carver, J.P. and Blout, E.R. (1969) J. Mol. Biol. 39, 307-313.

59. Brown, F.R., Dicorato, A., Lorenzi, G.P. and Blout, E.R. (1972) J. Mol. Biol. 63, 85-99.

60. Bruckner, P. and Eikenberry, E.F. (1984) Eur. J. Biochem, 140, 391-395.

61. Burstein, E.A., Vedenkina, N.S. and Ivkova, M.N. (1973) Photochemistry and Photobiology 18, 263-279.

62. Burgeson, R.E., El Adli, F.A., Kaitila, I.I. and Hollister, D.W. (1976) Proc. Natl. Acad. Sci. USA 73, 2579-2583.

63. Bush, C.A., Sarkar, B.K. and Kopple, K.D. (1978) Biochemistry 17, 4951-4955.

64. Butler, W.T. and Cunningham, L.W. (1966). J. Biol. Chem. 241, 3882-3888.

65. Bylund, D.B. and Krebs, E.A. (1975). J. Biol. Chem. 250, 6355-6363.

66. Cann, J.R., Stewart, J.M. and Matsueda, G.R. (1973) Biochemistry 12, 3780-3788.

67. Cantor, C.R. and Schimmel, P. (1980) *Biophysical Chemistry* Parts I and II, Freeman Company, San Franscisco.

68. Cardinale, G.J. and Udenfriend, S. (1974) Adv. in Enzymol. 41, 245-300.

69. Cardinale, G.J., Rhoads, R.E. and Udenfriend, S. (1971) Biochem. Biophys. Res. Commun. 43, 537-543.

70. Chang, C.T., Wu, C.S.C and Yang, J.T. (1978) Anal. Biochem. 91, 13-31.

71. Chen, Y.H. and Yang, J.T. (1971) Biochem. Biophys. Res. Commun. 44, 1285-1291.

72. Chen, Y.H., Yang, J.T., and Martinez, H.M. (1972) Biochemistry 11, 4120-4131.

73. Chen, Y.H., Yang, J.T. and Chan, K.H. (1974) Biochemistry 13, 3350-3359.

74. Chen-Kiang, S., Cardinale, G.J. and Udenfriend, S. (1977). Proc. Natl. Acad. Sci. USA 74, 4420-4424.

75. Chopra, R.K. and Ananthanarayanan, V.S. (1982) Proc. Natl. Acad. Sci. USA 79, 7180-7184.

76. Chou, P.Y. and Fasman, G.D. (1977) J. Mol. Biol. 115, 135-175.

77. Chrispeels, H.J. (1984) in *Methods in Enzymology* 107, 361-369.

78. Chung-Wong, R.S. and Bennick, A. (1980) J. Biol. Chem. 25, 5943-5948.

79. Citri, N. (1973) Adv. in Enzymol. 37, 397-648.

80. Cleland, W.W. (1964) Biochemistry 3, 480-482.

81. Comper, W. and Veis, A. (1977a) Biopolymers 16, 2113-2131.

82. Comper, W. and Veis, A. (1977b) Biopolymers 16, 2131-2142.

83. Cook, W.J., Einspann, H., Trapane, T.L., Urry, D.W. and Bugg, U.E. (1980). J. Am. Chem. Soc. 102, 5502-5505.

84. Cornish-Bowden, A. (1981) *Fundamentals of Enzyme Kinetics* p 75-85, Butterworths, London.

85. Counts, D.F., Cardinale, G.J. and Udenfriend, S. (1978) Proc. Natl. Acad. Sci. USA 75, 2145-2149.

86. Crawford, J.L., Lipscomb, W.N. and Schellman, C.G. (1973) Proc. Natl. Acad. Sci. USA 70, 538-542.

87. Crippen, G.M. and Yang, J.T. (1974) J. Phys. Chem. 78, 1127-1130.

88. Crisma, M., Fasman, G.D., Balaram, H. and Balaram, P. (1984) Int. J. Pept. Prot. Res. 23, 411-419.

89. Cunliffe, C.J. and Franklin, T.J., (1986) Biochem. J. 239, 311-315.

90. Deben, C.M. (1974) Macromolecules 7, 47-51.

91. De Jong, L. and Kemp, A. (1982) Biochim. Biophys. Acta 709, 142-145.
92. De Jong, L. and Kemp, A. (1984) Biochim. Biophys. Acta 787, 105-111.
93. De Jong, L., Albracht, S.P.J. and Kemp, A. (1982). Biochim. Biophys. Acta 704, 326-332.
94. Delaney, N.G. and Madison, V. (1982) J. Am. Chem. Soc. 104, 635-666.
95. De Waal, A., De Jong, L., Hartog, A.F. and Kemp, A. (1985) Biochemistry 24, 6493-6499.
96. De Waal, A., Hartog, A.F. and De Jong, L. (1987) Biochim. Biophys. Acta 912, 151-155.
97. Dickerson, R.E. and Geis, I. (1969) In *The Structure and Action of Proteins*, p 42. Harper and Row Publishers, New York..
98. Dixon, M. (1953) Biochem. J. 55, 161-171.
99. Doyle, B.B., Traub, W., Lorenzi, G.P. and Blout, E.R. (1971) Biochemistry 10, 3052-3060.
100. Doyle, B.B., Bendit, E.G. and Blout, E.R. (1975) Biopolymers 14, 937-957.
101. Drickamer, K., Dordal, M.S. and Reynolds, L. (1986) J. Biol. Chem. 261, 6878-6887.
102. Eastoe, J.E. (1967) In *Treatise on Collagen*, Vol. 1, 1-72. (Ed. Ramachandran, G.N.)
103. Edwards, C.A. and O'Brien, W.D. Jr. (1980) Clin. Chim. Acta. 104, 161-167.
104. Edman. J.C., Ellis, L., Blacher, R.W., Roth, R.A. and Rutter, W.J. (1985) Nature, 317, 267-273.
105. Engel, J. (1967) In *Conformation of Biopolymers* (Ed. Ramachandran, G.N.), Vol. 2, p 483-497, Academic Press, New York.
106. Engel, P.C. (1981) *Enzyme Kinetics - The Steady State Approach*.. Outline Studies in Biology. (1981) p 52, 2nd Edition. Chapman and Hall, London.
107. Engel, J., Kurtz, J., Katchalski, E. and Berger, A. (1966) J. Mol. Biol. 17, 255-272.
108. Engel, J., Chen, H.T., Prockop, D.J. and Klump, H. (1977) Biopolymers 16, 601-622.
109. Fasman, G.D. (1967) Poly α- amino acids, Dekker, New York.
110. Feramisco, J.R., Glass, D.B. and Krebs, E.G. (1980) J. Biol. Chem. 255, 4240-4245.
111. Fersht, A. (1983) In *Enzyme Structure and Mechanism*, p 85-101, W.H. Freeman and Company.
112. Fessler, L.I. and Fessler, J.H. (1974) J. Biol. Chem. 249, 7637-7646.
113. Fessler, J.H. and Fessler, L.I. (1978) Ann. Rev. Biochem. 47, 129-162.
114. Fietzek, P.P. and Kuhn, K. (1976). Int. Rev. Connect. Tissue. Res. 7, 1-60.

115. Fietzek, P.P., Wendt, P., Kell, I. and Kuhn, K. (1972a) FEBS Letters 26, 74-76.

116. Fietzek, P.P., Rexrodt, F.W., Wendt, P., Stark, M. and Kuhn, K. (1972) Eur. J. Biochem. 30, 163-168.

117. Fleischmajer, R.F., Olsen, B. and Kuhn, K. (Eds.) (1985) *Biology, Chemistry and Pathology of Collagen*, Annals of N.Y. Acad. Sci. Vol. 460.

118. Fraser, R.D.B. and McRae, T .P. (1973) *Conformation in Fibrous Proteins*, Academic Press, New York.

119. Freedman, R.B. and Hawkins, H.C. (1980) *The Enzymology of Posttranslational Modifications of Proteins*, Vol. 1.

120. Frey, M.H., Hexem, J. G., Leo, G.C., Tsang, P., Opella, S.J., Rockwell, A. and Gierasch, L.M. (1983) In *Peptides Structure and Function*, (Eds. Hruby, V.J. and Rich, D.J.), p 763-771.

121. Fujimoto, D. and Prockop, D.J. (1968) J. Biol. Chem. 243, 4138-4142.

122. Fujimoto, D. and Tamiya, N. (1962) Biochem. J. 84, 333-335.

123. Giam. C. Z., Chai, T., Hayashi, S. and Heury, W. (1984) J. Biochem. 141, 331-337.

124. Glass, D.B. and McPherson, M. (1986) J. Biol. Chem. 261, 5674-5679.

125. Gleason, F.K., Whittaker, M.M., Holmgren, A. and Jornvall, H. (1985) J. Biol. Chem., 260, 9567-9573.

126. Grant, M.E. and Prockop, D.J. (1972) N. Engl. J. Med. 286, 291-300.

127. Gray, W.R., Sandberg, L.B. and Foster, J.A. (1973) Nature 246, 461-466.

128. Greenfield, N. and Fasman, G.D. (1969) Biochemistry 8, 4108-4115.

129. Gross, J. (1976) In *Biochemistry of Collagen*, (Eds. Ramachandran, G.N. and Reddi, A.H.), p 275-310, Plenum Press, New York.

130. Gunzler, V., Hanauski-Abel, H.M., Myllyla, R., Mohr, J. and Kivirikko, K. (1987) Biochem. J. 242, 163-169.

131. Gustavson, K.H. (1955) Nature 175, 70-74.

132. Halme, J. and Jaaskelainen, M. (1970) Biochem. J. 116, 367-369.

133. Halme, J., Kivirikko, K.I. and Simons, K. (1970) Biochim. Biophys. Acta 198, 460-470.

134. Hanauski-Abel, H.M. and Günzler, V. (1982) J. Theor. Biol. 94, 421-455.

135. Harrington, W.F. and Von Hippel, P.H. (1961) Adv. Protein Chem. 16, 1-138.

136. Hayashi, O., Nozaki, M. and Abbot, M.T. (1975) In *The Enzymes*, (Ed. Boyer, P.D.) Vol. XII, p 119-189, Academic Press, New York.

137. Helbecque, N. and Loucheux-Lefebvre. M.H. (1978) Int. J. Pept. Prot. Res. 11, 353-362.

138. Helseth, D.L., Lechner, J.H. and Veis, A. (1979) Biopolymers 18, 3005-3014.

139. Hennessey, J.P. and Johnson, W.C. Jr. (1981) Biochemistry 20, 1085-1094.

149. Hobza, P., Hurych, J. and Zahradnik, R. (1973) Biochim. Biophys. Acta. 304, 466-472.

141. Hoffmann, H., Fietzek, P.P. and Kuhn, K. (1980) J. Mol. Biol. 141, 293-314.

142. Hollosi, M., Kawai, M. and Fasman, G.D. (1985) Biopolymers 24, 211-242.

143. Hoyhtya, M., Myllyla, R., Pjuva, J., Kivirikko, K.I. and Tryggvason, K. (1984) Eur. J. Biochem. 141, 477-482.

144. Hulmes, D.J.S., Miller, A., Parry, D.A.D., Piez, K.A. and Woodhead-Galloway, J. (1973) J. Mol. Biol. 79, 137-148.

145. Hurych, J. and Chvapil, M. (1965) Biochim. Biophys. Acta 97, 361-363.

146. Hutton, J.J., Kaplan, A. and Udenfriend, S. (1967a) Arch. Biochem. Biophys. 121, 384-391.

147. Hutton, J.J., Tappel, A.L. and Udenfriend, S. (1967b) Arch. Biochem. Biophys. 118, 231-340.

148. Hutton, J.J., Witkop, M.B., Kurtz, J., Berger, A. and Udenfriend, S. (1968) Arch. Biochem. Biophys. 125, 779-785.

149. Isemura, T., Okabayashi, H. and Sakakibara, S. (1968) Biopolymers 6, 307-332.

150. Jardetzky, O. (1980) Biochim. Biophys. Acta 621, 227-232.

151. Jencks, W.P. (1975) Adv. in Enzymol. 43, 219-410.

152. Jennes, D.D., Sprecher, C. and Johnson, W.C. (1976) Biopolymers 15, 513-521.

153. Jimenez, S., Harsch, M. and Rosenbloom, J. (1973) Biochem. Biophys. Res. Commun. 52, 106-114.

154. Josse, J. and Harrington, W.F. (1964) J. Mol. Biol. 9, 269-287.

155. Juva, K. and Prockop, D.J. (1969) J. Biol. Chem. 244, 6486-6492.

156. Kartha, G., Ashida, T. and Kakudo, M. (1974) Acta Cryst. B30, 1861-1866.

157. Katz, E.P. and Li, S.T. (1972) Biochem. Biophys. Res. Commun. 46, 1368-1373.

158. Kawai, M. and Fasman, G.D. (1978) J. Am. Chem. Soc. 100, 3630-3632.

159. Kedersha, N.L. and Berg, R.A. (1981) Collagen Rel. Res. 1, 345-353.

160. Kefalides, N.A. (1975) J. Invest. Dermatol. 65, 85-92.

161. Kemp, B.E., Bylund, D.B., Huang, T. and Krebs, E.G. (1975) Proc. Natl. Acad. Sci. USA 72, 3448-3452.

162. Kerr, S.J. and Borek, E. (1973) in *Enzymes*, 2nd Ed. 9, 167-195.

163. Kikuchi, Y., Fujimoto, D. and Tamiya, N. (1969) Biochem. J. 115, 569-574.

164. Kishida, Y., Berg, R.A., Sakakibara, S. and Prockop, D.J. (1976) quoted in Biochemistry of Collagen) p 181 (Eds. Ramachandran, G.N. and Reddi, A.H.), Plenum Press, New York.

165. Kivirikko, K.I. and Myllyla, R. (1979) Int. Rev. Connect. Tissue. Res. 8, 23-72.

166. Kivirikko, K.I. and Myllyla, R. (1980) In *The Enzymology of Posttranslational Modifications of Proteins*, (Eds. Freedman, R.B. and Hawkins, H.C.), Vol. 1, 53-104, Academic Press, London.

167. Kivirikko, K.I. and Myllyla, R. (1982) Methods in Enzymol. 82, 245-317.

168. Kivirikko, K.I. and Myllyla, R. (1984) In *Extracellular Matrix Biochemistry*, (Eds. Piez, K.A. and Reddi, A.H.), pp 83-118, Elsevier, New York.

169. Kivirikko, K.I. and Myllyla, R. (1985) In *Annals of the New York Academy of Sciences*, Vol. 460, pp 187-201. (Eds. Fleischmajer, R., Olsen, B.R. and Kuhn, K.)

170. Kivirikko, K.I. and Prockop, D.J. (1967a) J. Biol. Chem. 242, 4007-4012.

171. Kivirikko, K.I. and Prockop, D.J. (1967b) Arch. Biochem. Biophys. 118, 611-618.

172. Kivirikko, K.I. and Prockop, D.J. (1967c) Proc. Natl.Acad. Sci. USA 57, 782-789.

173. Kivirikko, K.I. and Prockop, D.J. (1967d) Biochem. J. 102, 432-442.

174. Kivirikko, K.I., Ganser, V., Engel, J. and Prockop, D.J. (1967) Hoppe-Seyler's Z. Physiol. Chem. 348, 1341-1344.

175. Kivirikko, K.I., Bright, H.J. and Prockop, D.J. (1968) Biochim. Biophys. Acta. 151, 558-567.

176. Kivirikko, K.I., Prockop, D.J., Lorenzi, G.P. and Blout, E.R. (1969) J. Biol. Chem. 244, 2755-2760.

177. Kivirikko, K.I., Suga, K., Kishida, Y., Sakakibara, S. and Prockop, D.J. (1971) Biochem. Biophys. Res. Commun. 45, 1591-1596.

178. Kivirikko, K.I., Kishida, Y., Sakakibara, S. and Prockop, D.J. (1972) Biochim. Biophys. Acta. 271, 347-356.

179. Kobayashi, Y., Sakai, R., Kakiuchi, K. and Isemura, T. (1970) Biopolymers 9, 415-425.

180. Kobayashi, Y., Suezaki, Y., Inonye, K., Tanaka, K., Kakiuchi, K. and Kyogoku, Y. (1977) Biopolymers 16, 247-275.

181. Koivu, J. and Myllyla, R. (1986) Biochemistry 25, 5982-5986.

182. Koivu, J., Myllyla, R., Helaakoski, T., Pihlajaniemi, T., Tasanen, K. and Kivirikko, K.I. (1987) J. Biol. Chem. 262, 6447-6449.

183. Kolaskar, A.S., Sarathy, K.P. and Sasisekharan, V. (1975) Curr. Sci. 44, 35-38.

184. Konev, S.V. (1967) *Fluorescence and Phosphorescence of Proteins and Nucleic Acids*. Plenum Press, New York.

185. Kopple, K.D., Go, A. and Pilipanskas, D.R. (1975) J. Am. Chem. Soc. 97, 6830-6838.

186. Koshland, D.E. Jr. (1976) Fed. Proc. 35, 2104-2111.

187. Kraut, J. (1977) Ann. Rev. Biochem. 46, 331-358.

188. Kresina, T.F. and Miller, E.J. (1979) Biochemistry 1, 3089-3097.

189. Kulonen, E. and Pikkarainen, J. (1970). In *Chemistry and Molecular Biology of the Intercellular Matrix*, Vol. 1, pp 81- 97. Academic Press, London.

190. Kuutti, E.R., Tuderman, L. and Kivirikko, K.I. (1975) Eur. J. Biochem. 57, 181-188.

191. Laemmli, U.K. (1970) Nature 227, 680-685.

192. Lapiere, C.M. and Nusgens, B. In *Biochemistry of Collagen* (Eds. Ramachandran, G. N. and Reddi, A. H.), pp 377-447, Plenum Press, New York.

193. Lazarev, Y.A., Grishkovsky, B.A. and Khromova, T.B. (1985) Biopolymers 24, 1449-1478.

194. Lee, E., Nemethy, G., Scheraga, H.A. and Ananthanarayanan, V.S. (1984a) Biopolymers 23, 1193-1206.

195. Lee, S., Mizuno, H., Nakamura, H., Kodera, Y., Kato, T., Go, N. and Izumiya, N. (1984b) FEBS Letters 174, 310-313.

196. Lewis, R.N., Momany, F.A. and Scheraga, H.A. (1973) Biochim. Biophys. Acta 303, 211-229.

197. Lineweaver, H. and Burk, D. (1934) J. Amer. Chem. Soc. 56, 658-666.

198. Lowry, O., Rosebrough, N., Farr, A. and Randall, R. (1951) Anal. Chem. 193, 265275.

199. Lwebuga-Mukasa, J.S., Lappi, S. and Taylor, P. (1976) Biochemistry 15, 1425-1434.

200. Madison, V. (1974) In *Peptides, Polypeptides and Proteins* (Eds. Blout, E. R., Bovey, F. A., Goodman, M. and Lotan, S.), pp 89-98, John Wiley, New York.

201. Madison, V. and Kopple, K.D. (1980) J. Am Chem. Soc. 102, 4855-4863.

202. Madison, V. and Schellman, J. (1970a) Biopolymers 9, 511-567.

203. Madison, V. and Schellman, J. (1970b) Biopolymers 9, 569-588.

204. Madison, V. and Schellman, J. (1972) Biopolymers 11, 1041-1076.

205. Majamaa, K., Hanauski-Abel, H.M., Gunzler, V. and Kivirikko, K. (1984) Eur. J. Biochem. 138, 339-245.

206. Majamaa, K., Turpeenniemi-Hujanen, T.M., Latipaa, P., Gunzler, V., Hanauski-Abel, H.M., Hassiners, I.E. and Kivirikko, K.I. (1985) Biochem. J. 229, 127-133.

207. Majamaa, K., Gunzler, V., Hanauski-Abel, H.M., Myllyla, R. and Kivirikko, K. (1986) J. Biol. Chem. 261, 7819-7823.

208. Mandel, R. and Holzworth, G. (1973) Biopolymers 12, 655-674.

209. Matsuzaki, T. (1974) Acta Cryst. B30, 1029-1036.

210. Mattice, W.L. and Mandelkern, L. (1970) Macromolecules 3, 199-201.

211. Mattice, W.L. and Mandelkern, L. (1971) Biochemistry 10, 1926-1933.

212. Maxfield, F.R., Leach, S.J., Stimpson, E.R., Powers, S.P. and Scheraga, H.A. (1979) Biopolymers 18, 2507-2521.

213. McGee, J.O.D., Rhoads, R.E. and Udenfriend, S. (1971) Arch. Biochem. Biophys. 144, 343-351.

214. McClure, W.R and Jovin, M. (1975) J. Biol.Chem. 250, 4073-4080.

215. Meggio, F., Donella Deana, A. and Pinna, L.A. (1981) Biochim. Biophys. Acta 662, 1-7.

216. Miller, E.J. (1985) In *Chemistry, Biology and Pathology of Collagen*, Annals of N.Y. Academy of Sci. (Eds. Fleischmajer, R., Olsen, B.R. and Kuhn, K.), 460, 1-13.

217. Miyazawa, T. (1967) in *Poly-α-amino acids*, (Ed. Fasman, G.D.), pp 70-103, Marcel Dekker, New York.

218. Muller-Eberhard, H.J. (1975) Ann. Rev. Biochem. 44, 697-724.

219. Murphy, L. and Rosenbloom, J. (1973) Biochem. J. 135, 249-251.

220. Myllyla, R. and Seppa, H. (1979) Biochem. J. 182, 311-316.

221. Myllyla, R., Risteli, L. and Kivirikko, K.I. (1975) Eur. J. Biochem. 58, 517-521

222. Myllyla, R., Tuderman, L. and Kivirikko, K.I. (1977) Eur. J. Biochem. 80, 349-357.

223. Myllyla, R., Kuulti-Savolainen, E.R. and Kivirikko, K.I. (1978) Biochem. Biophys. Res. Commun. 83, 441-448.

224. Myllyla, R., Schubotz, L.M., Weser, U. and Kivirikko, K.I. (1979) Biochem. Biophys. Res. Commun. 89, 98-102.

225. Myllyla, R., Majamaa, K., Gunzler, V., Hanauski-Abel, H.M. and Kivirikko, K.I. (1984) J. Biol. Chem. 259, 5403-5405.

226. Nemethy, G. and Printz, M. (1972) Macromolecules 5, 755-758.

227. Nietfield, J.J. and Kemp, A. (1980) Biochim. Biophys. Acta 613, 349-358.

228. Nietfield, J.J. and Kemp, A. (1981) Biochim. Biophys. Acta 657, 159-167.

229. Nietfield, J.J., De Jong, L. and Kemp, A. (1982) Biochim. Biophys. Acta 704, 321-325.

230. Nimni, M.E. and Deshmukh, K. (1973) Science 181, 751-752.

231. Okabayashi, H., Isemura, T. and Sakakibara, S. (1968) Biopolymers 6, 323-330.

232. Okuyama, K., Tanaka, N., Ashida, T., Kakudo, M., Sakakibara, S. and Kishida, Y. (1972) J. Mol. Biol., 72, 571-576.

233. Okuyama, K., Tanaka, N., Ashida, T. and Kakudo, M. (1976) Bull. Chem. Soc. Japan 49, 1805- *.

234. Olsen, R.B., Berg, R.A., Kivirikko, K.I. and Prockop, D.J. (1973) Eur. J. Biochem. 35, 135-147.

235. Paik, W.K. and Kim, S. (1975) Adv. Enzymol. 42, 227-286.

236. Pankalainen, M., Aro, H., Simons, K. and Kivirikko, K.I. (1970) Biochim. Biophys. Acta 221, 559-565.

237. Park, E., Tanzer, M.L. and Church, R.L. (1975) Biochem. Biophys. Res. Commun. 63, 1-10.

238. Paul, P.K.C. and Ramakrishnan, C. (1985) J. Biomol. Struct. Dynam. 2, 879-895.

239. Pease, L.G., Niu, C.A. and Zimmerman, G. (1979) J. Amer. Chem. Soc 100, 1279-1286.

240. Pease, L.G., Frey, M.H. and Opella, S.J. (1981) J. Amer. Chem. Soc. 103, 467-468.

241. Peterkofsky, B. and Diblasio, R. (1973) Anal. Biochem. 66, 279-286.

242. Piez, K.A. (1976) In *Biochemistry of Collagen* (Eds. Ramachandran, G.N. and Reddi, A.H.), p 1-44, Plenum Press, New York.

243. Pihlajaniemi, T., Helaakoski, T., Tasanen, K., Myllyla, R., Huhtala, M.L., Koivu, J. and Kivirikko, K.I. (1987) EMBO J. 6, 643-649.

244. Plesner, I.W. (1986) Biochem. J. 239, 175-178.

245. Popenoe, E.A., Aransan, R.B. and Van Slyke, D.D. (1969) Arch. Biochem. Biophys. 133, 286-292.

246. Porter, R.R. and Reid, K.B.M. (1978) Nature 275, 699-704.

247. Prockop, D.J., Kaplan, A. and Udenfriend, S. (1963) Arch. Biochem. Biophys. 101, 499-503.

248. Prockop, D.J. and Juva, K. (1965a) Biochem. Biophys. Res. Commun. 18, 54-59.

249. Prockop, D.J. and Juva, K. (1965b) Proc. Natl. Acad. Sci. USA 53, 661-668.

250. Prockop, D.J. and Kivirikko, K.I. (1969) J. Biol. Chem. 244, 4838-4842.

251. Prockop, D.J., Juva, K. and Engel, J. (1967) Z. Physiol. Chem. 348, 553-560.

252. Prockop, D.J., Berg, R.A., Kivirikko, K.I. and Uitto, J. (1976) In *Biochemistry of Collagen* (Eds. Ramachandran, G.N. and Reddi, H.A.), p 163-273, Plenum Press, New York.

253. Provencher, S.W. and Glockner, J. (1981) Biochemistry 20, 33-37.

254. Ramachandran, G.N. , (1967) *Treatise on Collagen* (Ed. Ramachandran, G.N.) Academic Press, London.

255. Ramachandran, G.N. and Chandrasekharan, R. (1968) Biopolymers 6, 1649-1658.

256. Ramachandran, G.N. and Kartha, G. (1954) Nature 174, 269-270.

257. Ramachandran, G.N. and Kartha, G. (1955a) Nature 176, 593-595.

258. Ramachandran, G.N. and Kartha, G. (1955b) Proc. Indian. Acad. Sci. A45, 215- *.

259. Ramachandran, G. N. and Ramakrishnan, C. (1976) In *Biochemistry of Collagen* (Eds. Ramachandran, G.N. and Reddi, A.H.), p 45-84, Plenum Press, New York.

260. Ramachnadran, G.N. and Reddy, A.H. (1976) *Biochemistry of Collagen* (Eds. Ramachandran, G.N. and Reddi, A.H.) Plenum Press, New York.

261. Ramachandran, G.N., Doyle, B.B. and Blout, E.R. (1968) Biopolymers 6, 1771-1775.

262. Ramachandran, G.N., Bansal, M. and Bhatnagar, R.S. (1973) Biochim. Biophys. Acta 322, 166-171.

263. Ramachandran, G.N., Lakshminarayanan, A.V. and Kolaskar, A.S. (1973) Biochim. Biophys. Acta 303, 8-13.

264. Rao, C.N.R. (1963) *Chemical Applications of Infrared Spectroscopy* Academic Press, N.Y. and London.

265. Rao, N.V. and Adams, E. (1978) J. Biol. Chem. 253, 6327-6330.

266. Rao, C.P., Nagaraj, R., Rao, C.N.R. and Balaram, P. (1980) Biochemistry 19, 425-431.

267. Rao, C.P., Balaram, P. and Rao, C.N.R. (1983) Biopolymers 22, 2091-2104.

268. Rapaka, R.S., Sorenson, K.R., Lee, S.D. and Bhatnagar, R.S. (1976) Biochim. Biophys. Acta. 429, 63-71.

269. Rapaka, R.S., Renugopalakrishnan, V., Urry, D.W. and Bhatnagar, R.S. (1978) Biochemistry 17, 2892-2898.

270. Reed, L. J. (1974) Acc. Chem. Res. 7, 40-46.

271. Reed, L.L. and Johnson, P.L. (1973) J. Amer. Chem. Soc. 95, 7523-7524.

272. Renugopalakrishnan, V., Kluomann, P.H.B. and Bhatnagar, R.S. (1984) Biopolymers 23, 623-627.

273. Reid, K.B.M. (1977) Biochem. J. 161, 247-251.

274. Reid, K.B.M. and Porter, R.R. (1976) Biochem. J. 155, 19-23.

275. Reid, K.B.M. and Thompson, O.P. (1978) Biochem. J. 173, 863-868.

276. Rhoads, R.E. and Udenfriend, S. (1968) Proc. Natl. Acad.Sci. USA 60, 1473-1478.

277. Rhoads, R.E. and Udenfriend, S. (1969) Arch. Biochem. Biophys. 133, 108-111.

278. Rhoads, R.E. and Udenfriend, S. (1970) Arch. Biochem. Biophys. 139, 329-339.

279. Rhoads, R.E., Udenfriend, S. and Bornstein, P. (1971) J. Biol. Chem. 246, 4138-4142.

280. Risteli, J., Tuderman, L. and Kivirikko, K.I. (1976) Biochem. J. 158, 369-376.

281. Rose, D., Gierasch, L.M. and Smith, J.A. (1985) Adv. Protein Chem. 37, 1-109.

282. Rosenberg, T.L. and Richardson, J.M. (1977) Biochemistry 16, 3550-3558.

283. Sadava, D. and Chrispeels, H.J. (1971) Biochim. Biophys. Acta 227, 278-287.

284. Sakakibara, S., Kishida, Y., Kikuchi, Y., Sakai, R. and Kakiuchi, K. (1968) Bull. Chem. Soc. Japan 41, 1173 - *.

285. Sakakibara, S., Kishida, Y., Okayama, K., Tanaka, N., Ashida, T. and Kakudo, M. (1972) J. Mol. Biol. 65, 371-373.

286. Sakakibara, S., Inonyi, K., Shudo, K., Kishida, Y., Kobayashi, Y. and Prockop, D.J. (1973) Biochim. Biophys. Acta 303, 198-202.

287. Salem, G. and Traub, W. (1975) FEBS Letters 51, 94-99.

288. Sandberg, L.B. (1976) Int. Rev. Connect. Tissue Res. 7, 159-210.

289. Sasisekharan, V. (1959a) Acta. Cryst. 12. 897-903.

290. Sasisekharan, V. (1959b) Acta. Cryst. 12, 903-909.

291. Saxena, V.P. and Wetlauffer, D.B. (1971) Proc. Natl. Acad. Sci. USA 68, 969-972.

292. Scatturin, A., Tamburro, A.M., Delpra, A. and Bordignon, E. (1975) Int. J. Pept. Protein Res. 7, 425-435.

293. Schimmel, P.R. and Flory, P.J. (1968) J. Mol. Biol. 34, 105-120.

294. Schlifpenbacher, R., Wenzel, S., Lottspeich, F. and Sumper, M. (1986) FEBS Letters 209, 57-62.

295. Schulz, G.E. and Schirmer, R.H. (1979) (Ed. C.R. Cantor) *Principles of Protein Structure* Springer-Verlag, New YorK.

296. Shenolikar, S. and Cohen, P. (1978) FEBS Letters 86, 92-98.

297. Sherman, L.A. and Gefter, M.L. (1976) J. Mol. Biol. 103, 61-76.

298. Siegel, R.C. (1979) Int. Rev. Connect. Tissue. Res. 8, 73-118.

299. Smith, J.A. and Pease, L.G. (1980) CRC, Crit. Rev. Biochem. 8, 315-399.

300. Spiro, R.G. and Spiro, M.J. (1971) J. Biol. Chem. 246, 4910-4918.

301. Stern, A., Gibbons, W.A. and Craig, L.C. (1968) Proc. Natl. Acad. Sci. USA 61, 734-741.

302. Stimson, E.R., Zimmerman, S.S. and Scheraga, H.A. (1977) Macromolecules 10, 1049-1060.

303. Stolz, M., Furthmayer, H. and Timpl, R. (1973) Biochim. Biophys. Acta 310, 461-468.

304. Subramanian, E. and Lalitha, V. (1983) Biopolymers 22, 833-838.

305. Sutoh, K. and Noda, H. (1974a) Biopolymers 13, 2391-2404.

306. Sutoh, K. and Noda, H. (1974b) Biopolymers 13, 2461-2475.

307. Suzuki, F. and Koyama, E. (1969) Biochim. Biophys. Acta 177, 154-156.

308. Takahashi, K., Kang, Y.K., Nemethy, G., Scheraga, H.A. and Ananthanarayanan, V.S. (1987) Biopolymers 26, 1781-1788.

309. Tanaka, M. and Shibata, H. (1985) Eur.J. Biochem 151, 291-297.

310. Tanaka, I., Kozima, T. and Ashida, T. (1977) Acta Cryst. B33, 116-119.

311. Tanaka, I. and Ashida, T., Shimonishi, Y. and Kakudo, M. (1979) Acta Cryst. B35, 110-114.

312. Tanzer, M.L., Church, R.L., Yaeger, J.A., Wampler, D.E. and Park, E.D. (1974) Proc. Natl. Acad. Sci. USA 71, 3009-3010.

313. Tanzer, M.L. (1976) In *Biochemistry of Collagen*, 137-162. (Eds. Ramachandran, G.N. and Reddi, A.H.) Plenum Press, New York.

314. Tiffany, M.L. and Krimm, S. (1969a) Biopolymers 8, 347-359.

315. Tiffany, M.L. and Krimm, S. (1969b) Biopolymers 8, 695-697.

316. Tiffany, M.L. and Krimm, S. (1972) Biopolymers 11, 2309-2316.

317. Tinker, D.A., Krebs, E.G., Feltham, I., Attah-Poku, S.K.A.P. and Ananthanarayanan, V.S. (1986) Fed. Proc., 45, 1799.

318. Torchia, D.A. and Bovey, F.A. (1971) Macromolecules 4, 246-251.

319. Torchia, D.A. and Lyerla, J.R. Jr. (1974) Biopolymers 13, 97-114.

320. Torchia, D.A., Dicorato, A., Wong, S.C.K., Deber, C.M. and Blout, E.R. (1972) J. Amer. Chem. Soc. 94, 609-615.

321. Tryggvason, K., Risteli, J. and Kivirikko, K.I. (1978) Clin. Chim. Acta 82, 233-240.

322. Tuderman, L., Kuutti, E.R. and Kivirikko, K.I. (1975) Eur. J. Biochem. 52, 9-16.

323. Tuderman, L., Myllyla, R. and Kivirikko, K.I. (1977a) Eur. J. Biochem. 80, 341-348.

324. Tuderman, L., Oikarinen, A. and Kivirikko, K.I. (1977b) Eur. J. Biochem. 78, 547-556.

325. Urry, D.W., Cunningham, W.D. and Ohinishi, T. (1974a) Biochemistry 13, 609-616.

326. Urry, D.W., Long, M.M., Ohnishi, T. and Jacobs, M. (1974b) Biochem. Biophys. Res. Commun 61, 1427-1433.

327. Urry, D.W. and Ohinishi, T. (1974) In *Peptides, Polypeptides and Proteins* (Eds. Blout, E.R., Bovey, F.A., Goodman, M and Lotan, S.), pp 230-247.

328. Urry, D.W., Venkatachalam, C.M., Long, M.M. and Prasad, K.U. (1983) in *Conformation in Biology*, The Ferstchrift (eds. Srinivasan, R. and Sharma, R.H.) pp 11-27, Adenine Press, New York.

329. Urivetzky, M., Frei, J.M. and Meilman, E. (1966) Arch. Biochem. Biophys. 177, 224-231.

330. Venkatachalam, C.M. (1968) Biopolymers 6, 1425-1436.

331. Venkatachalapathi, Y.V. and Balaram, P. (1979) Nature 281, 83-84.

332. Venkatachalapathi, Y.V. and Balaram, P. (1981) Biopolymers 20, 1137-1145.

333. Vistiça, D.T., Ahrens, F.A. and Ellison, W.R. (1977) Arch. Biochem. Biophys. 179, 15-23.

334. Vogel, A.I. (1957) In *A Text-Book of Practical Organic Chemistry* pp 163-17, Longmans, Green and Co., London.

335. Woody, R.W. (1974) In *Peptides, Polypeptides and Proteins*, (Eds. Blout, E.R., Bovey, F. A, Goodman, M and Lotan, S.), pp 339-3349.

336. Yagi, Y., Tanaka,I., Yamane, T. and Ashida, T. (1983) J. Amer. Chem. Soc. 105, 1242-1246.

337. Yamada, Y., Mudryi, M. and De Crombruggue, B. (1983) J. Biol.Chem 258, 14914-14919.

338. Yonath, A. and Traub, W. (1969) J. Mol. Biol. 43, 461-477.

339. Yee, R.Y., Englander, S.W. and Von Hoppel, P.M. (1974) J. Mol. Biol. 83, 1-16.

340. Zijenah, L.S. and Ananthanarayanan, V.S. (1987) in the Proceedings of the Can. Fed. Biol. Soc. Winnipeg (Manitoba), PP 177 (Abstract).

341. Zimmerman, S.S. and Scheraga, H.A. (1977) Biopolymers 16, 811-843.

* Original references could not be obtained.

Appendix A

Secondary Structure Analysis of Proline-containing Peptides from CD Data

It is now commonly accepted that the CD spectra of peptides and proteins represent the average of contributions from the constituent structural elements (Chang *et al.*, 1978). Using the reference CD spectra obtained from the polypeptides or from proteins of known structures for α-helix, β-sheet and unordered structures, it is possible to calculate the individual proportions of these three structural elements from the experimental CD spectrum of a protein (Greenfield and Fasman, 1969; Saxena and Wetlauffer, 1971; Chen and Yang, 1971; Chen *et al.*, 1972; 1974). More recently, investigators have incorporated the contributions from another important secondary structure namely, the β-turn in their analysis (Chang *et al.*, 1978; Brahms and Brahms, 1980; Provencher and Glockner, 1981; Hennessey and Johnson, 1981). As discussed in Chapter 4, the major conformational features of Pro-containing peptides have been found to be the β-turn in non-polar solvents and an extended PP-II like conformation in a polar medium like water. In this Appendix, conformations of the Pro-containing peptides are analyzed in terms of mixtures of β-turn, PP-II and random structures. A similar, but limited analysis has been carried out by Aubert *et al.* (1975) in the case of the Pro-rich protein (PRP) from human parotid saliva.

A.1. Reference CD Spectra for β-turn, PP-II and Random-coil Conformations

A.1.1. CD Spectrum of β-turn

The vacuum-UV CD spectrum of N-Acetyl-Pro-Gly-Leu-OH in TFE at low temperature was shown to be the representative of an isolated β-turn (Brahmachari *et al.*, 1979). The CD spectral features of this tripeptide include the presence of a relatively strong negative band at about 182 nm, a strong positive band at about 202 nm and another weak negative band around 225 nm. The mean residue ellipticities of the bands were found to be about -1,650 deg.cm^{2}dmol^{-1} at 225 nm and +9,000 deg.cm^{2}dmol^{-1} at 202 nm. These will be translated into the following molar ellipticity values of -5,000 deg.cm^{2}dmol^{-1} at 225 nm and +14,000 deg.cm^{2}dmol^{-1} at 202 nm at 25 °C (Brahmachari *et al.*, 1982). These authors have also shown that tripeptides of the type N-Acetyl-Pro-Gly-X-OH (where X = Gly, Ala, Leu, Val, Ile or Phe) adopt the β-turn conformation in order-promoting non-polar solvents (Brahmachari *et al.*, 1982). Qualitatively; these features are very similar to the class B spectrum calculated for types I and II β-turns by Woody (1974) in tripeptides. However, the magnitudes at the extrema seem to be quite different. Woody's calculated CD spectrum for type I β-turn exhibits about -5,000 deg.cm^{2}dmol^{-1} around 225 nm and about +29,000 deg.cm^{2}dmol^{-1} around 206 nm. On the other hand, the CD spectrum for type II β-turn exhibits about -2,500 deg.cm^{2}dmol^{-1} around 227 nm and about +30,000 deg.cm^{2}dmol^{-1} around 208 nm (Woody, 1974).

Brahms *et al.* (1977) reported a vacuum-UV CD spectrum for β-turn in

$(Ala_2\text{-}Gly_2)_n$ which exhibits -5,200 deg.cm^{2}dmol^{-1} at 227 nm and +63,000 deg.cm^{2}dmol^{-1} at 207 nm. The magnitudes at the extrema in the case of $(Ala_2\text{-}Gly_2)$ are very high when compared to the N-Acetyl-Pro-Gly-Leu-OH (Brahmachari *et al.*, 1982) or Woody's calculated spectra (Woody, 1974).

Urry *et al.* (1974a) reported CD spectra for cyclic peptide models of elastin. The CD spectrum of HCO-(Val-Pro-Gly-Gly)$_n$-Val-OMe (where n = 40) in TFE at -10 °C was reported to represent that of a typical β-turn which exhibits a negative band at 223 nm (mean residue ellipticity, -2,400 deg.cm^{2}dmol^{-1}) and a positive band at 203 nm (mean residue ellipticity, +2,300 deg.cm^{2}dmol^{-1}. The evidence for the presence of β-turn in this peptide was also obtained from proton magnetic resonance (PMR) studies. However, the magnitude of the ellipticity values of this compound seems to be very low when compared to all the above examples. In view of the mean residue ellipticity values obtained in the case of the test peptides used in the present studies, for example, t-Boc-Val-Pro-Gly-Val-OH and t-Boc-Gly-Val-Pro-Gly-Val-OH in TFE, the choice of the CD spectrum of N-Acetyl-Pro-Gly-Leu-OH as the reference CD spectrum seems to be more reasonable. The ratio of the positive and negative ellipticity values for this peptide in the β-turn conformation in TFE agrees well with that expected from Woody's calculations (Brahmachari *et al.*, 1979). Therefore, the CD spectrum of N-Acetyl-Pro-Gly-Leu-OH in TFE was chosen as the reference CD spectrum for β-turn conformation which is presented in Figure A-1.

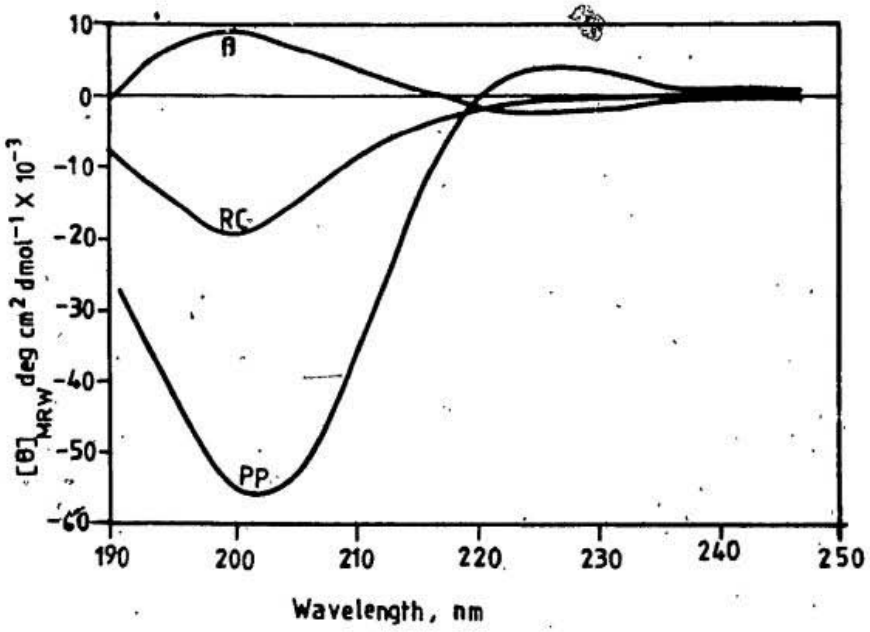


Figure A-1: Reference CD Spectra

Reference CD spectra for β-turn, PP-II and random-coil structures. The β-turn spectrum is obtained from N-Ac-PGL-OH in TFE at 0 °C; PP-II spectrum is obtained from poly(Pro) (M_r of 6,000) in TFE also at 0 °C; Random-coil spectrum is obtained from $(Pro\text{-}Pro\text{-}Gly)_5$ in 4 M $CaCl_2$ at 25 °C. The concentration of the peptides is 2 mg/ml ; cell size = 0.01 cm; θ is expressed as mean residue ellipticity.

A.1.2. CD Spectrum of PP-II Structure

The CD spectrum of poly(Pro) of M_r of 6,000 in TFE at low temperature exhibits a positive band around 225 nm and a very strong negative band at about 200 nm characterizing the type II conformation (Figure A.1). This spectrum is very similar to that reported by Tiffany and Krimm, (1969 a,b, 1972); Mandel and Holzworth (1973); Jennes *et al.*, 1976). The molar residue ellipticity of the positive band is about $+4{,}000 \pm 200$ deg.cm^{2}dmol^{-1} and that of the negative band is about $-56{,}000 \pm 1{,}200$ deg.cm^{2}dmol^{-1}. The spectrum in water at room temperature is very similar to that in TFE except that the θ_{225} values are decreased to about -3,000 deg.cm^{2}dmol^{-1} and θ_{200} to about -49,000 deg.cm^{2}dmol^{-1}. This indicates that the PP-II structure, unlike the β-turn, is not much destabilized by polar solvents. The CD spectrum in TFE and at low temperature was used as the reference for PP-II structure in the present analysis. Since most of the peptides used here are tetrapeptides, comparison with tetraproline (Pro_4) could be expected to give a better-fit between the computed and experimental spectra. However, CD spectra of $(Pro)_4$ in TFE or other non-polar solvents indicated the possible presence of the cis-peptide bonds (data not shown). Other investigators also indicated the possibility of cis-peptide bonds in small oligo-prolines (Isemura *et al.*, 1968; Okabayashi *et al.*, 1968; Deber, 1974). Therefore, the choice of PP-II (of higher molecular weight) which gives a characteristic all trans PP-II spectrum, seems to be more reasonable when compared to $(Pro)_4$. The reference CD spectrum for PP-II conformation is presented in Figure A.1.

A.1.3. CD Spectrum of Unordered or Random-coil Structure

Polypeptides in the presence of high concentrations of electrolytes such as $CaCl_2$ have been shown to exist in unordered structures that give CD spectra different from those of "extended-helical structures" of polyamino acids such as poly(Lys) or poly(Glu) in water at neutral pH (Tiffany and Krimm, 1969 a,b; Mattice and Mandelkern, 1971). These CD spectra are characterized by a strong negative band around 200 nm. It is known and should be noted here, that polypeptides give qualitatively similar spectra when they are in unordered conformation, although the magnitudes of the negative band at lower wavelengths vary from peptide to peptide (Jennes *et al.*, 1976; Tiffany and Krimm, 1969a,b). Since most of the peptides studied here have Pro and Gly residues as constituent amino acids, a polypeptide containing these residues, namely, $(Pro\text{-}Pro\text{-}Gly)_5$ has been chosen as the reference compound. The CD spectrum of this polypeptide in 4M $CaCl_2$ exhibits a single strong negative band at about 200 nm with a mean residue ellipticity of $-20{,}000 \pm 700$ deg.cm^{2}dmol^{-1}. This spectrum is very similar to that of unordered structure reported for poly(Pro), poly(Glu) and poly(Lys) (Tiffany and Krimm, 1969 a,b; Mattice and Mandelkern, 1971) and other polypeptides (Doyle *et al.*, 1971; Scatturin *et al.*, 1975). The reference CD spectrum for unordered structure used in the present analysis is shown in Figure A-1.

A.2. Method of Analysis

Using the above reference CD spectra for β-turn, PP-II and unordered structures, several series of spectra were generated for various mixtures of these conformations. The generation of these spectra was based on the assumption that the CD of a Pro-containing peptide at each wavelength can be expressed as a linear combination of the contributions from the β-turn, PP-II and unordered structures. A similar approach was used by Saxena and Wetlauffer (1971; Chen *et al.* (1972) and other investigators in the determination of secondary structures from the CD spectra of proteins. This can be expressed as follows :

$$[\Theta]_{\lambda}^{Obs} = f_{\beta t}[\Theta]_{\lambda}^{\beta t} + f_{PPII}[\Theta]_{\lambda}^{PPII} + f_{rc}[\Theta]_{\lambda}^{rc}$$

where $[\Theta]^{\beta t}$, $[\Theta]^{PP\text{-}II}$ and $[\Theta]^{rc}$ are the molar ellipticity contributions from $f_{\beta t}$, $f_{PP\text{-}II}$ and f_{rc}, the fractions of β-turn, PP-II and random-coil structures, respectively, at a given wavelength. Such sets of equations were applied in the wavelength range of 250-190 nm, to obtain the observed CD spectrum in terms of the sum of the intrinsic molar ellipticities of the three structures in different fixed proportions. During this treatment, the sum of fractions of all the structural components was fixed to be unity and the fraction of a particular conformation was always positive or zero i.e. ($\sum_1^i f_i = 1$ and $f_i \geq 0$)

Several series of CD spectra were thus generated each representing a unique combination of β-turn, PP-II and unordered structures. These CD spectra served as a library of standard spectra against which the test peptides were compared so

as to obtain the possible combination of conformations that may be present in these test peptides.

The normalized standard deviations between the experimental and computed spectra were calculated using the following formula :

$$S.D = \sqrt{\frac{\sum_{1}^{n}(\theta_{\lambda,n}^{Obs} - \theta_{\lambda,n}^{Cal})}{n-1}}$$

Where

$[\theta]_{\lambda,n}^{Obs}$ = Ellipticity in deg.cm^{2}dmol^{-1} at the given wavelength (λ, n) observed experimentally

$[\theta]^{Cal}$ = Ellipticity in deg.cm^{2}dmol^{-1} at λ, n in the computed (calculated) spectrum

n = Number of data points (i-e number of wavelengths)

n−1 = Degrees of freedom

A.3. Results

The following general observations may be made on the effect of the individual structural components on the computed CD spectrum.

1. The positive band around 200 nm (due to π-π^*) transition of the peptide amide) may be a sensitive marker for the β-turn and even slight changes in the relative proportions of this conformation may be monitored using this band.

2. The destabilization of PP-II structure by the unordered structure decreases the magnitude of the negative band at 200 nm much more than the positive band at 225 nm of the PP-II helix.

3. The admixture of the β-turn has nearly equal influence on the positive and negative bands of the PP-II structure. On the other hand, since the unordered structure does not have any significant contribution around 225 nm but does around 200 nm, the development of the unordered structure seems to affect mainly the 200 nm negative band of the PP-II structure. Therefore, the changes observed at 225 nm are interpretable conveniently in terms of the PP-II and β-turn conformers.

4. The negative band of PP-II spectrum around 200 nm seems to be dependent on the length of the peptide chain (as judged by the comparison of the CD s of higher and lower molecular weight polyprolines).

5. Based on considerations (2), (3) and (4), the positive band may be a better indicator for the PP-II conformation than the negative band.

6. In summary, the introduction of unordered structures in the β-turn or PP-II structures decreases the magnitudes of the positive and negative CD bands of these ordered structures, while their general spectral features are preserved.

A.3.1. Determination of Secondary Structures of the Peptides from their CD Spectra

The analysis was carried out by comparing the experimental CD spectrum of a given peptide in a given solvent with the computed spectra shown in Figures A-2 to A-9. In addition to the over-all fit of the experimental spectrum in comparison with that of the synthetic spectrum, the values at the wavelength maxima were given specific importance. Since the synthetic spectra were

generated in steps of 10% variations of one or more conformations at each time, it is not surprising to see that the experimental spectra did not always match exactly with the computed spectra. In those instances, an intermediate value of the two possible computed was tentatively assigned to the experimental spectrum in terms of the percentages (f_β, $f_{PP\text{-}II}$ and f_{rc}) of different mixtures of the conformers. After this initial coarse curve-fitting, fine-tuning was carried out in order to obtain better curve-fitting. In addition to the above manual analysis, the analysis was attempted using a computer program. While the values given by the computer are very similar to those obtained by manual analysis in some cases, the computer curve-fitting was totally unacceptable in other cases, especially in view of the original spectral features of these peptides. The reason for this apparent discrepancy is not understood and attempts to improve the program are underway.

The values of f_β, f_{PPII} and f_{rc} of different peptides obtained using the above data analysis were presented and discussed in Chapter 5 (Table 5.5). In this Appendix, the results are presented in Figures A-10 to A-26 (placed at the end of this Appendix).

A general examination of these results indicate that :

1. Pro-containing peptides are capable of existing in an equilibrium mixture of multiple conformations, namely, β-turn, PP-II and unordered structures and

2. The nature of the solvent determines the different proportions of each of these conformations. As discussed in Chapter 5, in the organic solvent TFE, the

folded, intra-molecularly H-bonded β-turn conformation is expresed more; on the other hand, in aqueous medium, the extended conformation of PP-II becomes predominant. Thus, a given peptide would have larger proportions of the β-turn and smaller amounts of PP-II in TFE, while in water the reverse would be the case.

A.4. Discussion

The method of analysis using CD spectra of known conformations as reference spectra has been successfully applied earlier to proteins (Greenfield and Fasman, 1969; Saxena and Wetlaufer, 1971; Chen *et al.*, 1972; Chang *et al.*, 1978; Brahms and Brahms, 1980). The analysis done here is slightly different, in that it deals with relatively small oligopeptides which have not hitherto been analyzed in this way. However, the technique used is essentially the same. The major consideration in using this technique is the choice of the reference CD spectra. Unlike the case with the protein analysis, very little is known about the reference states in small peptides. The reason for this is that small linear peptides have an enormous conformational flexibility in solution. Thus, it is difficult to obtain a single conformation for a peptide (in solution) whose CD can be used as the reference. In the present analysis, the reference CD spectra have been chosen rather empirically for reasons discussed below :

Variations in the CD spectra of oligo- and polyprolines have been observed based on the molecular weight and chainlength of the peptides. Although, the individual proline residues in these oligo- and polymers have their ϕ and ψ angles fixed appropriately for the PP-II like extended conformation, Okabayashi *et al.*

(1968) and Deber (1974) showed from CD, IR and NMR that the characteristic appearance of the helical conformation starts with $n \geq 4$ (i.e. 4 proline residues). X-ray studies on t-Boc-$(Pro)_3$ by Kartha *et al.* (1974) and on t-Boc-$(Pro)_4$ by Matsuzaki (1974) indicated that they adopt the PP-II like extended structure in solid-state. However, Helbecque and Louheux-Lefebebvre (1978) showed by CD and NMR that while Gly-$(Pro)_4$ exhibits a CD spectrum similar to that of PP-II (a negative band at about 225 nm), Gly-$(Pro)_3$ does not give a CD spectrum typical of PP-II. In the present analysis also marked effect of chainlength on the CD spectra (especially on the lower wavelength negative bands) was noticed. Therefore, there is a certain degree of uncertainty about the reference spectrum for 100% PP-II conformation.

The reference peptide for β-turn, namely, N-Acetyl-Pro-Gly-Leu-OH has been studied by vacuum-UV CD studies (Brahmachari *et al.*, 1979) and has been found to be qualitatively very similar to the class B spectrum proposed by Woody (1974) for type I and II β-turns. It is also similar to that of $(Val\text{-}Pro\text{-}Gly\text{-}Gly)_n$ studied by Urry *et al.* (1974a). However, X-ray analysis has not been done on this peptide. The presence of type II β-turn has been implied in this peptide by indirect evidence that an analogous compound, N-Acetyl-Pro-Gly-Phe-OH adopts type II β-turn in solid state (Brahmachari *et al.*, 1979). Therefore, again there is some assumption about the 100% β-turn structure.

There is considerable disagreement about the 100% unordered CD spectrum which seems to differ from peptide to peptide, protein to protein and peptides to proteins, except for the presence of a negative band around 200 nm (Tiffany and

Krimm, 1969 a,b; Chen *et al.*, 1972; Mattice and Mandelkern, 1971; Jennes *et al.*, 1976). In the absence of a typical CD spectrum for 100% unordered conformation, other investigators also had to make certain assumptions about the reference spectrum for 100% unordered structure in their conformational analyses (Greenfield and Fasman, 1969; Chen *et al.*, 1972 and Brahms and Brahms, 1980). The fact that, in the present analysis, only peptide conformations have been used as references to compare and analyze another set of peptides, should make these assumptions less objectionable in contrast to the assumptions made in extrapolating the polypeptide data to proteins.

As in the case of any structural investigation, the ultimate verification for this structural analysis will be obtained only from X-ray analysis of these peptides. Again, an assumption is made that the conformations of peptides would be the same in both solid state and in solution despite the fact that different stabilization forces operate in different media.

In the cases of Boc-VPGV-OH and Boc-GVPGV-OH, no PP-II type structure was detected by this CD analysis. This is in agreement with the studies on Boc-VPGV-OH by CD and IR (Chapter 5) and by X-ray diffraction (Yagi *et al.*, 1983) which indicate that the extended structure present may be similar to the β-strand, rather than the PP-II type structure. Since only the β-turn, PP-II and random-coil structures are considered for the reference CD spectra, the present analysis would not be able to give information on the nature of the extended structure present in these peptides. Consideration of the CD spectra of other extended structures like β-sheet structure may improve the conformational analysis on these peptides.

In spite of all these limitations, the method of conformational analysis presented here seems to provide a reasonable means of determining the different proportions of conformers presented by the peptide molecules in solution.

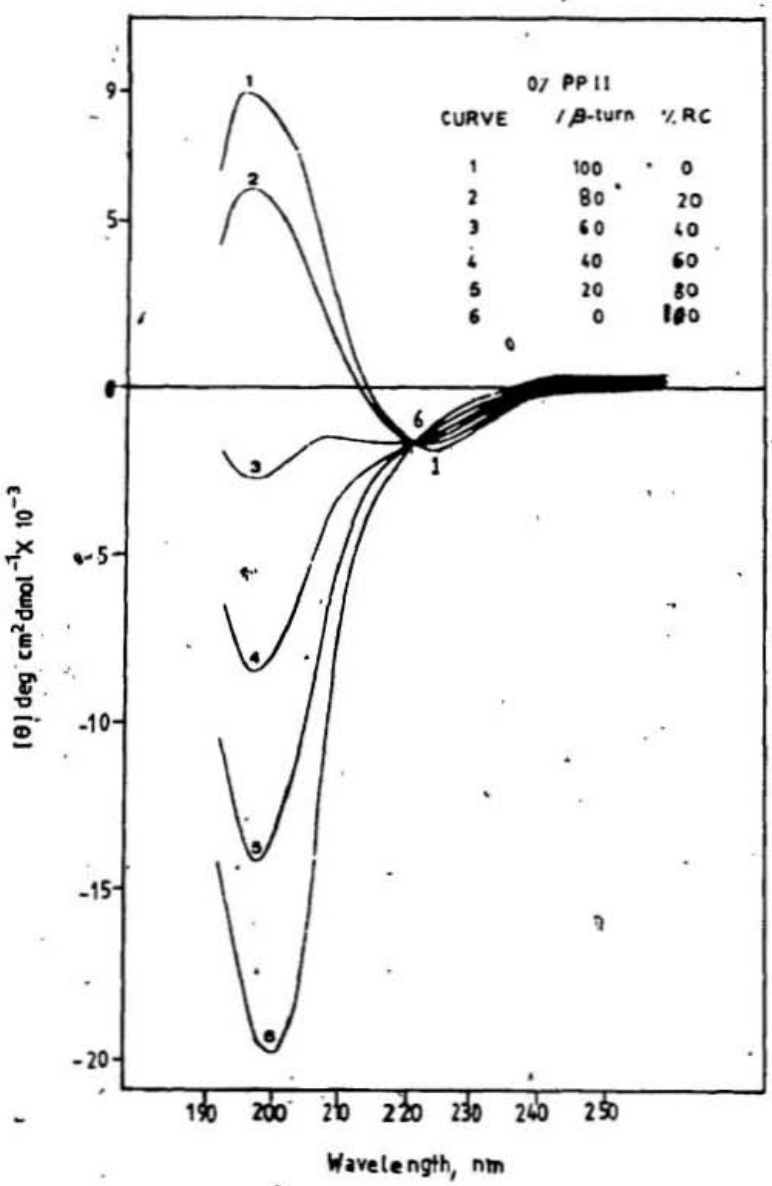


Figure A-2: Combination Spectra for 0% PP-II + 0-100% β-turn or Random-coil Structures

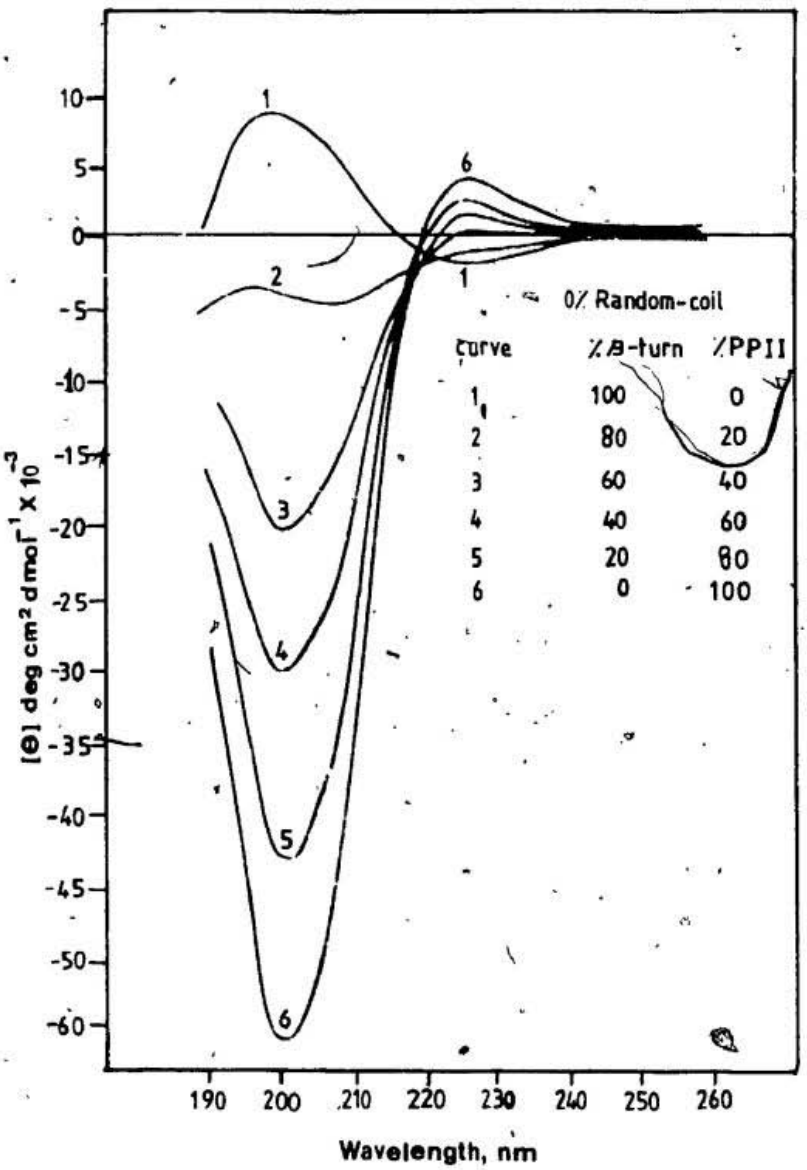


Figure A-3: Combination Spectra for 0% Random-coil + 0-100% β-turn or PP-II Structures

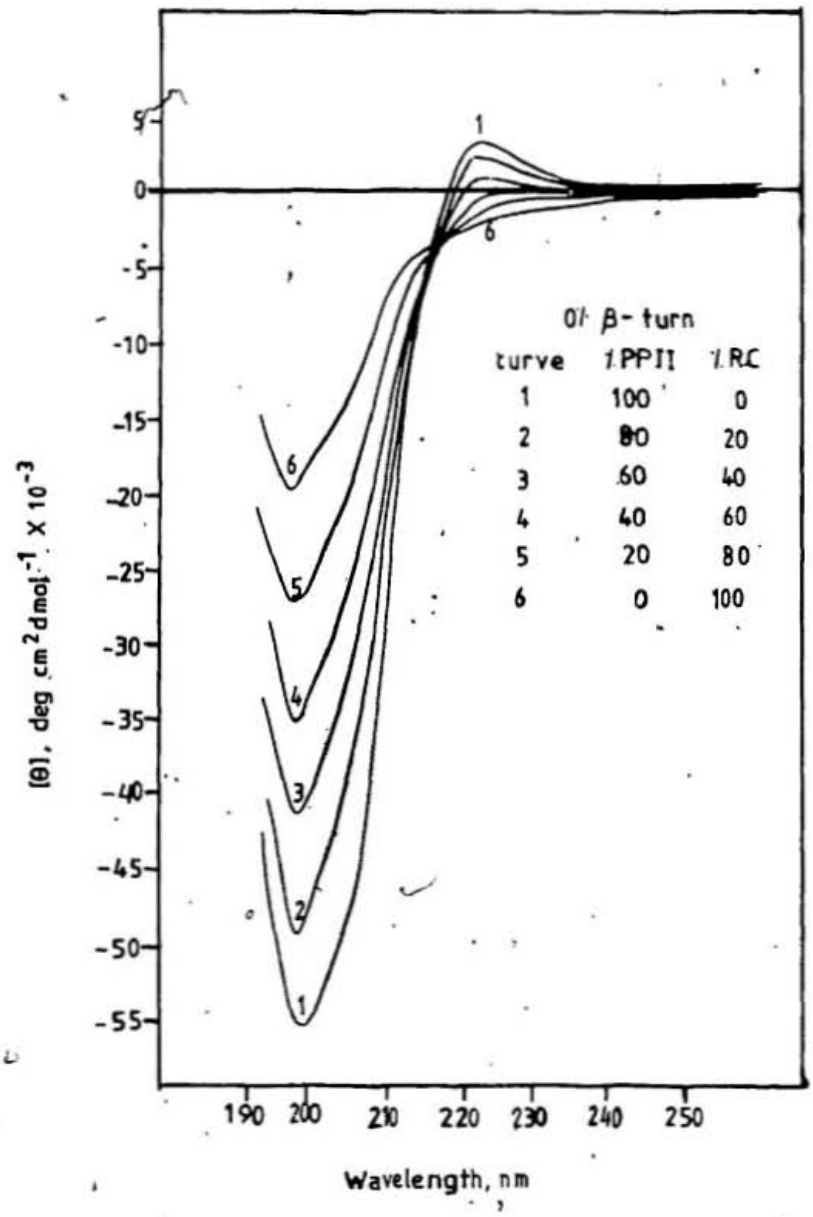


Figure A-4: Combination Spectra for 0% β-turn + 0-100% PP-II or Random-coil Structures

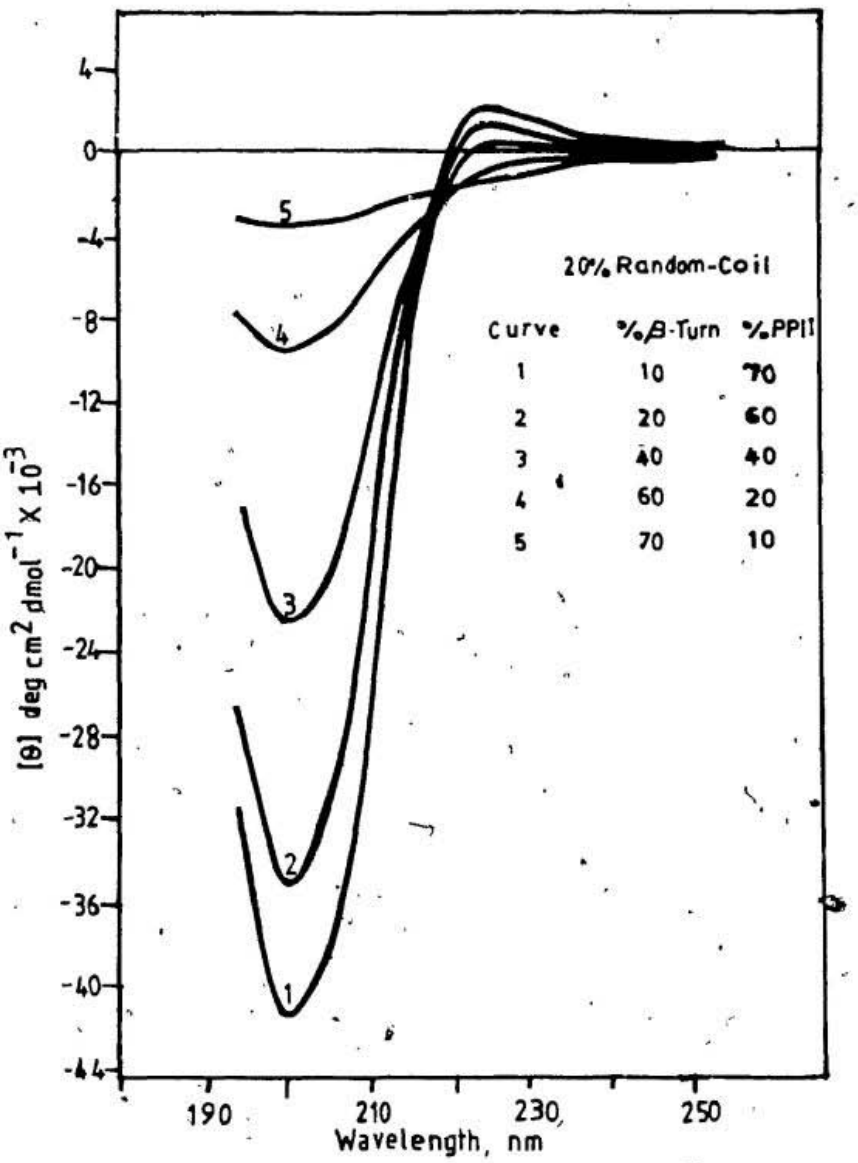


Figure A-5: Combination Spectra for 20% Random-coil + 10-70% PP-II or β-turn Structures

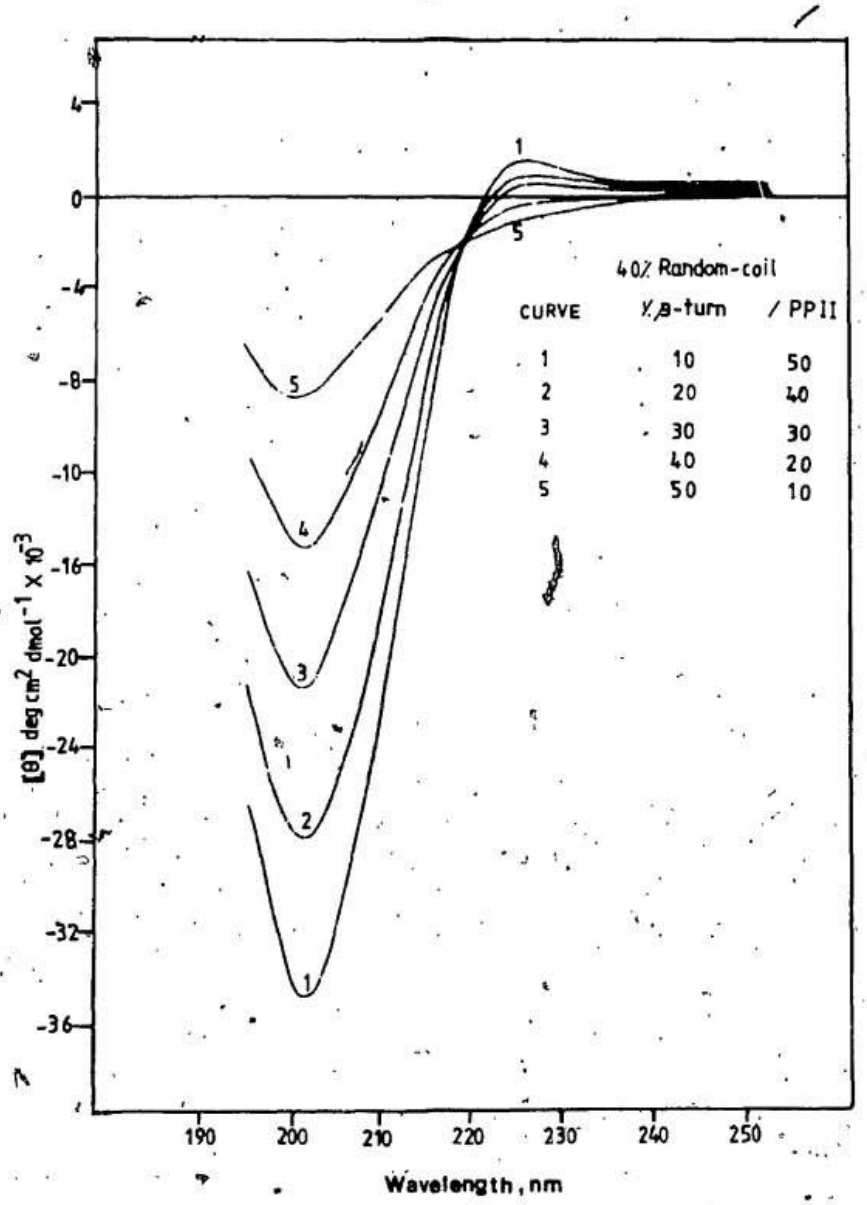


Figure A-6: Combination Spectra for 40% Random-coil + 10-50% PP-II or β-turn Structures

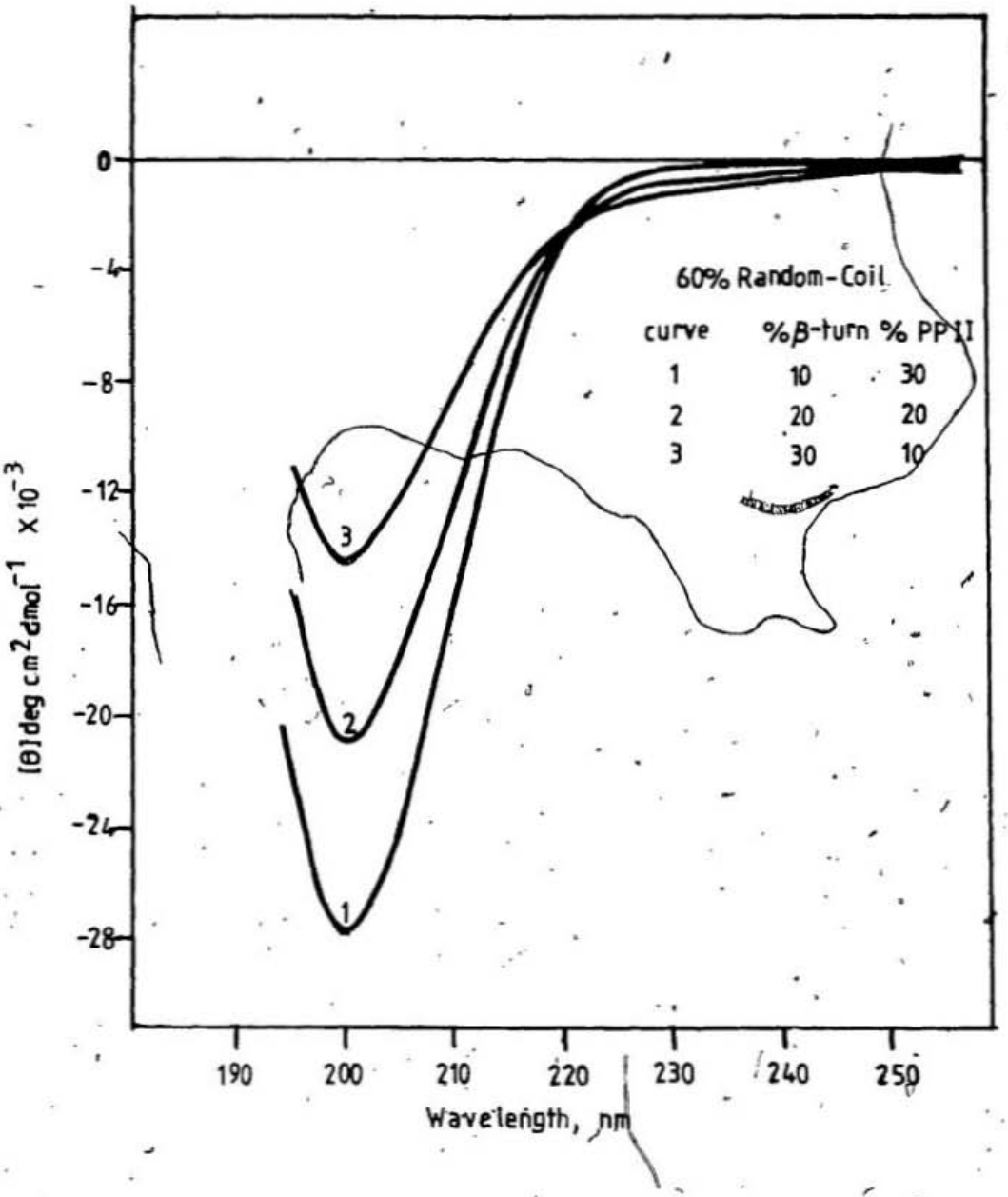


Figure A-7; Combination Spectra for 60% Random-coil + 10-30% PP-II or β-turn Structures

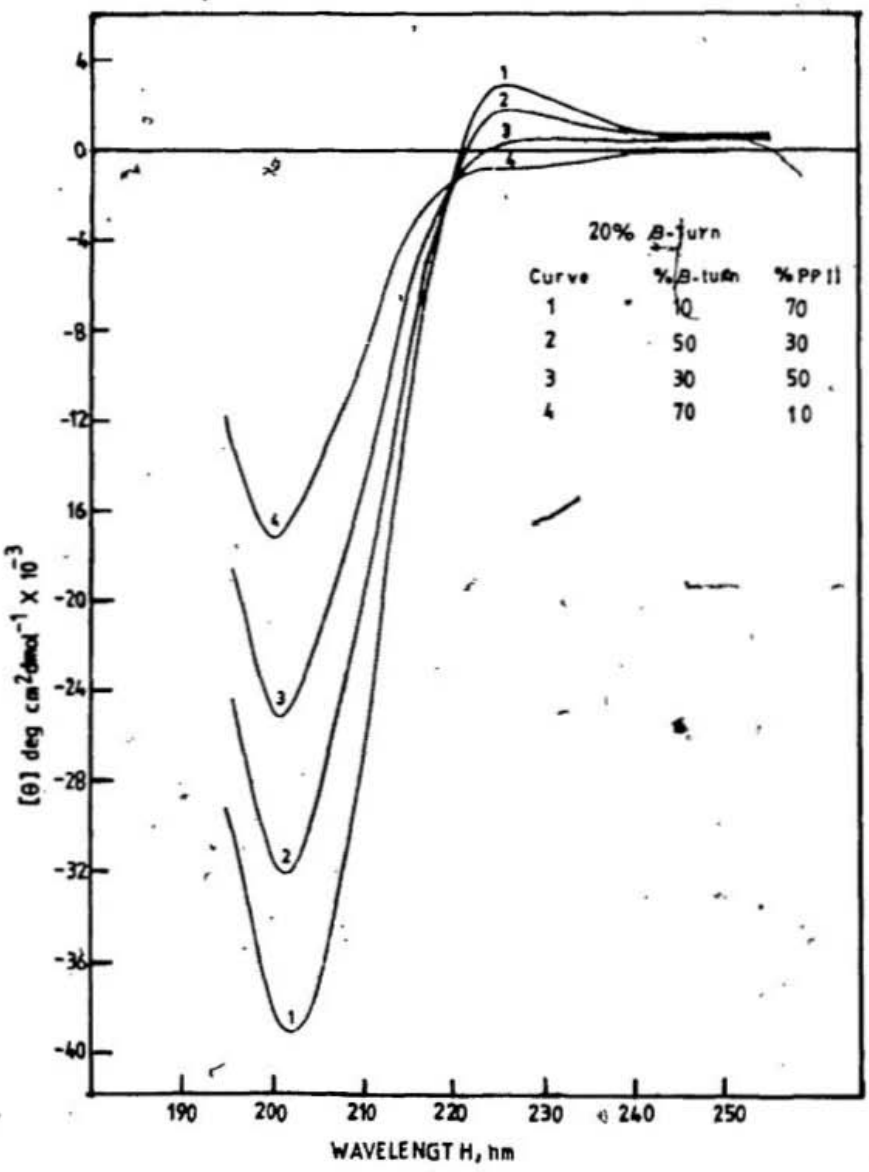


Figure A-8: Combination Spectra for 20% β-turn + 10-70% Random-coil or PP-II Structures

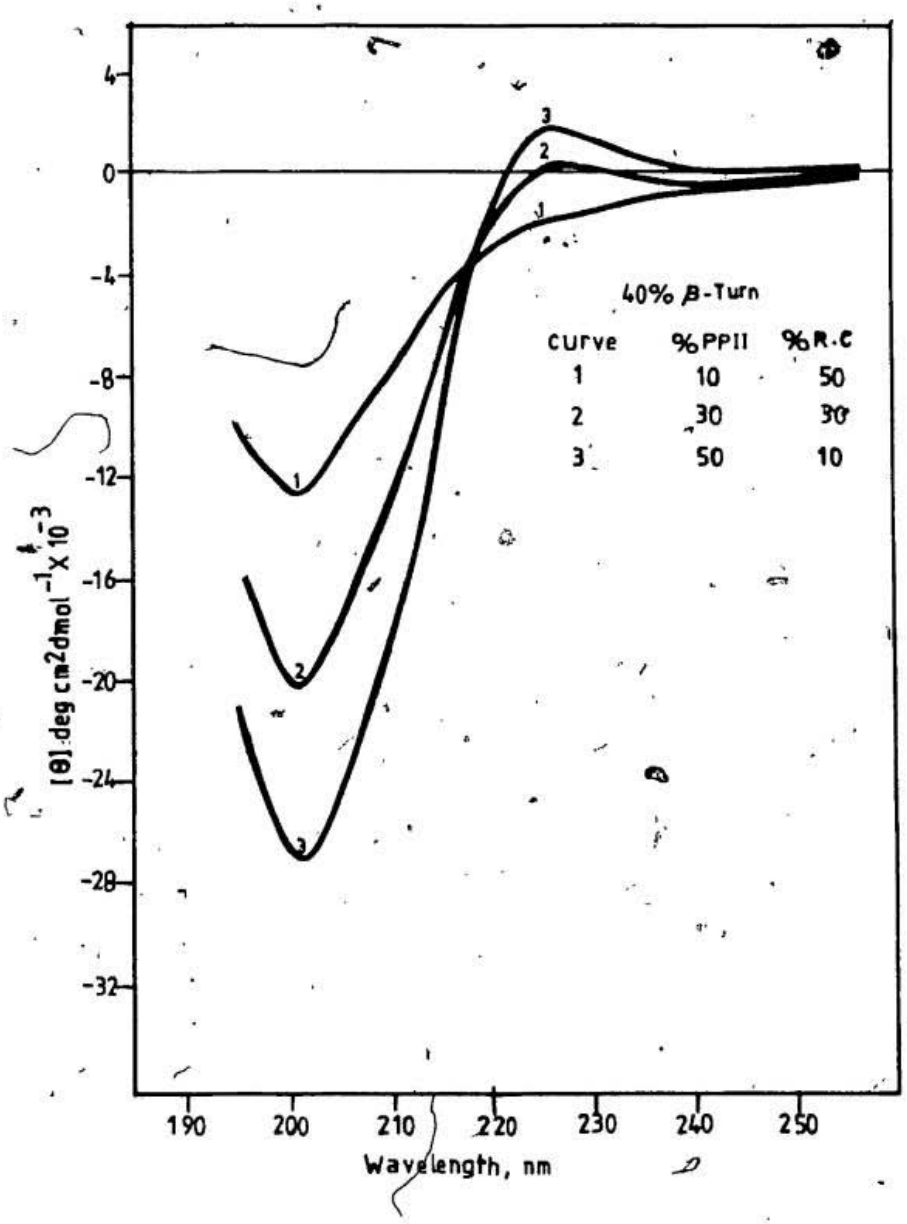


Figure A-9: Combination Spectra for 40% β-turn + 10-50% Random-coil or PP-II Structures

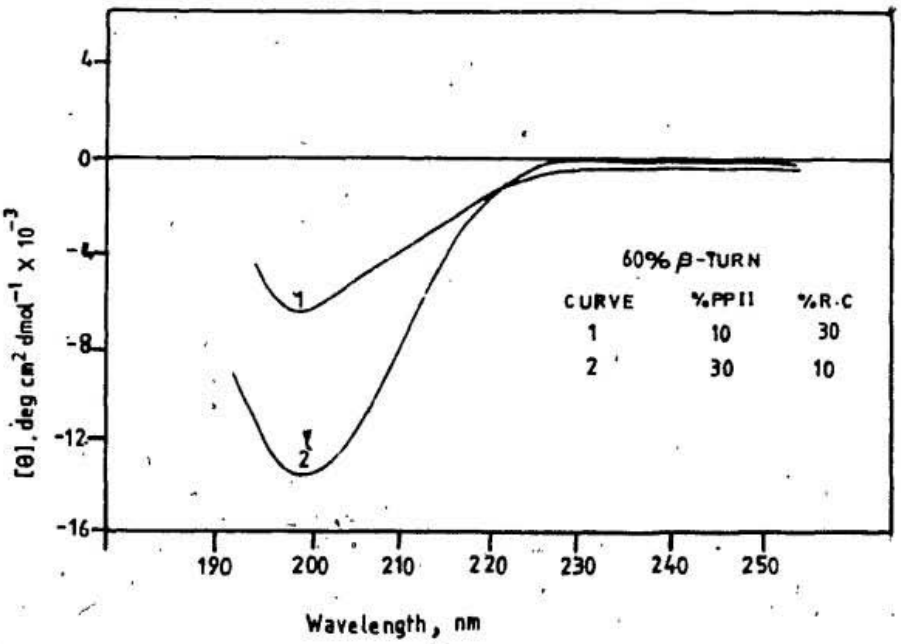


Figure A-10: Combination Spectra for 60% β-turn + 10-30% Random-coil or PP-II Structures

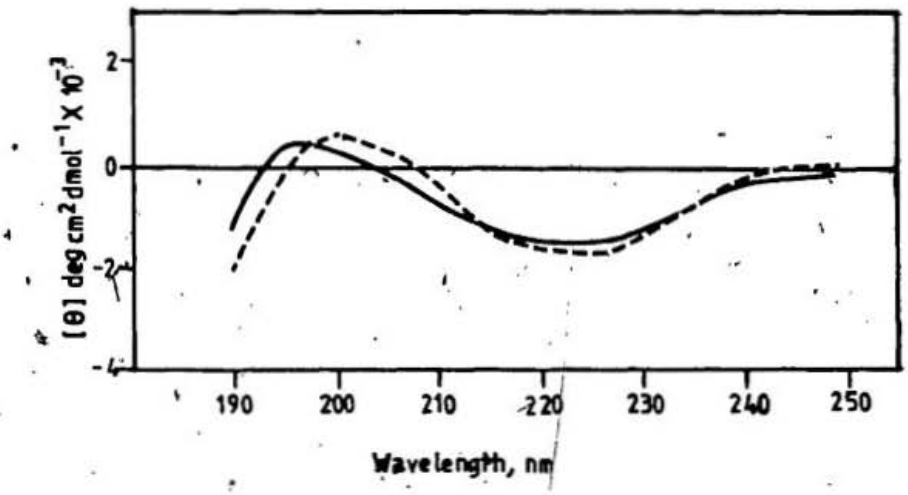


Figure A-11: CD Analysis of Boc-PPGNHCH$_3$ in TFE

(—) Experimental spectrum; (- - -) Computed spectrum for

70% β-turn + 30% random-coil structures.

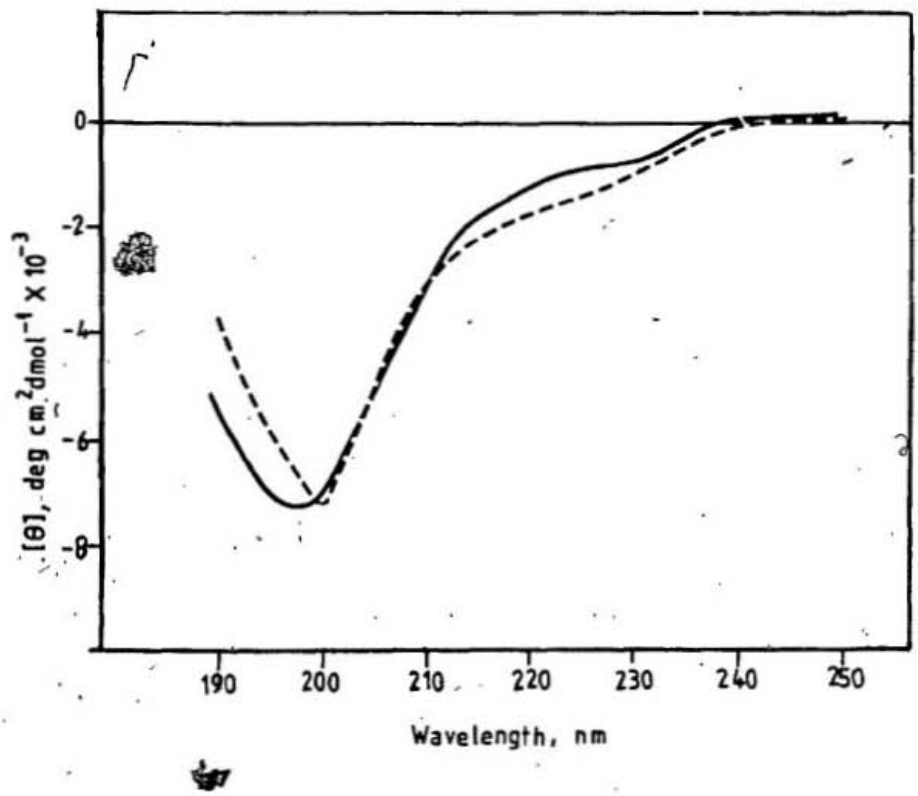


Figure A-12: CD Analysis of Boc-PPGP-OH in TFE

(—) Experimental spectrum; (- - -) Computed spectrum for

45% β-turn + 55% random-coil structures.

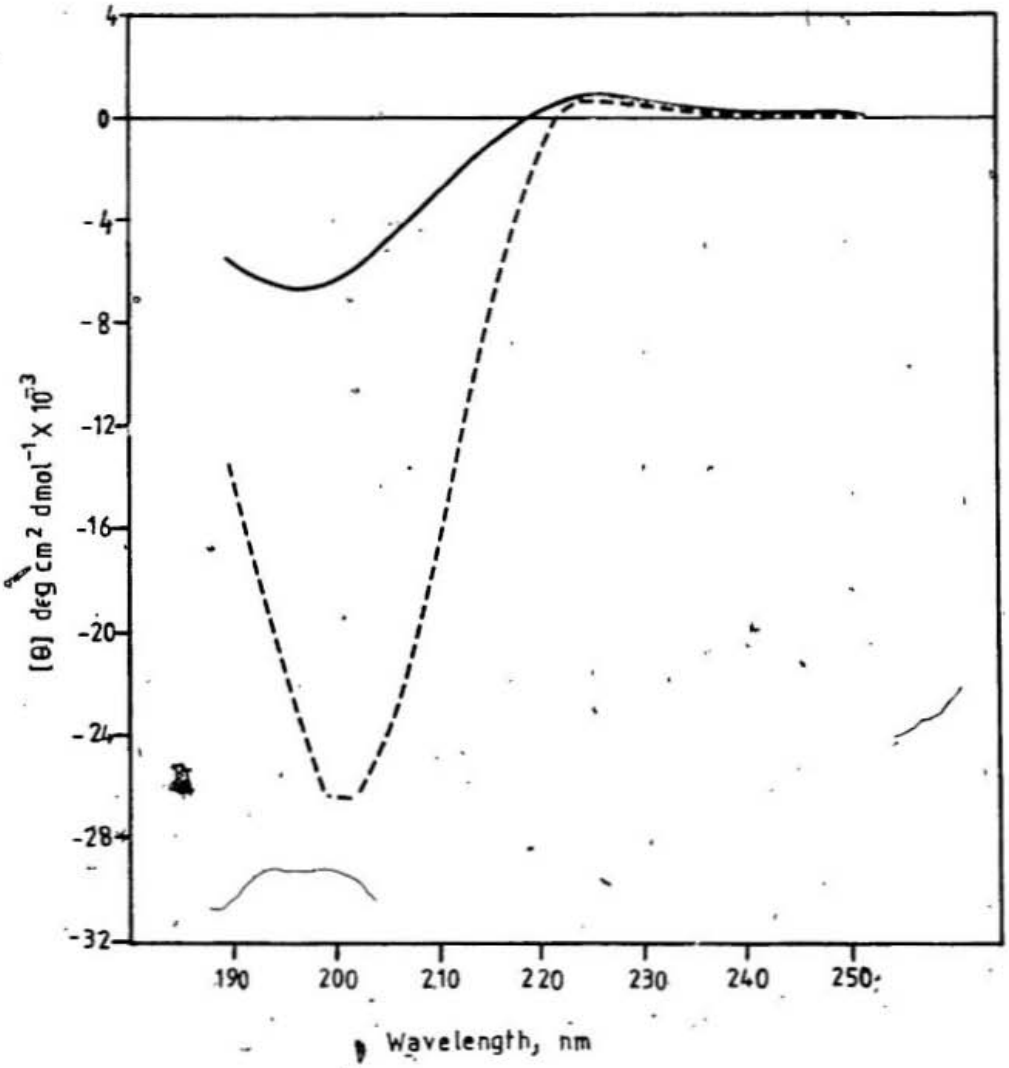


Figure A-13: CD Analysis of Boc-PPGP-OH in Water

(—) Experimental spectrum; (- - -) Computed spectrum for

35% β-turn + 45% PP-II + 20% random-coil structures.

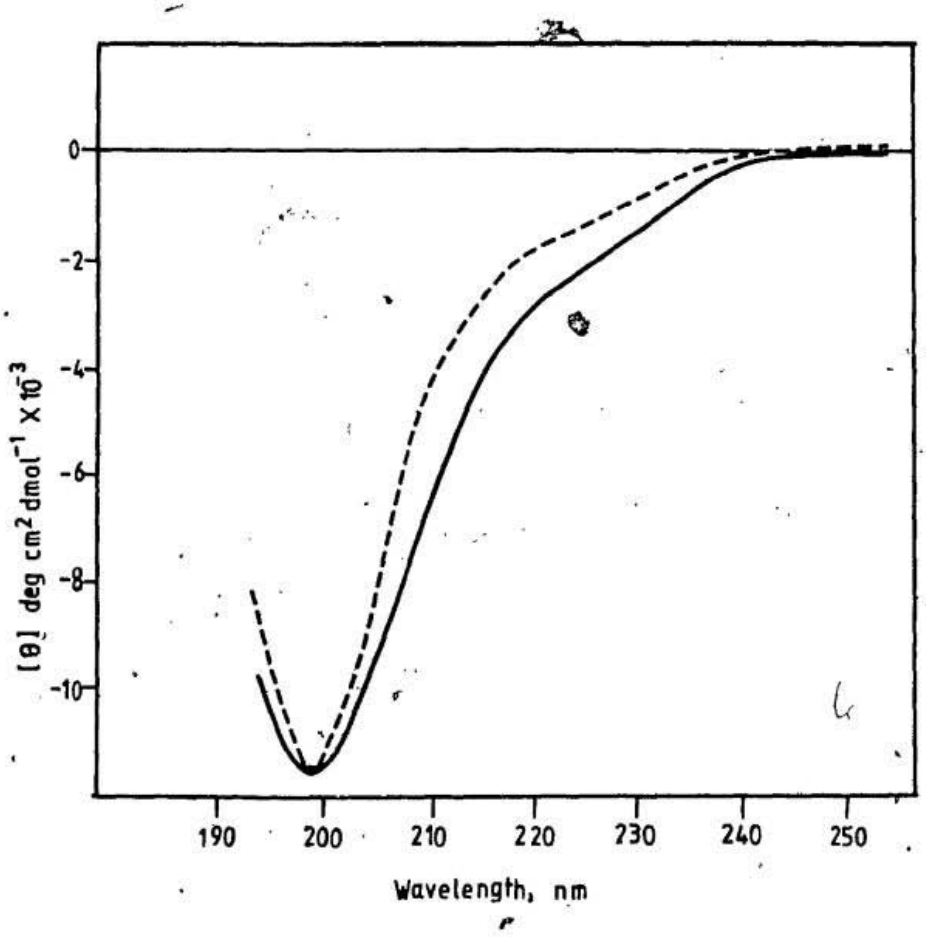


Figure A-14: CD Analysis of Boc-PPGP-$NHCH_3$ in TFE

(—) Experimental spectrum; (- - -) Computed spectrum for

30% β-turn + 70% random-coil structures.

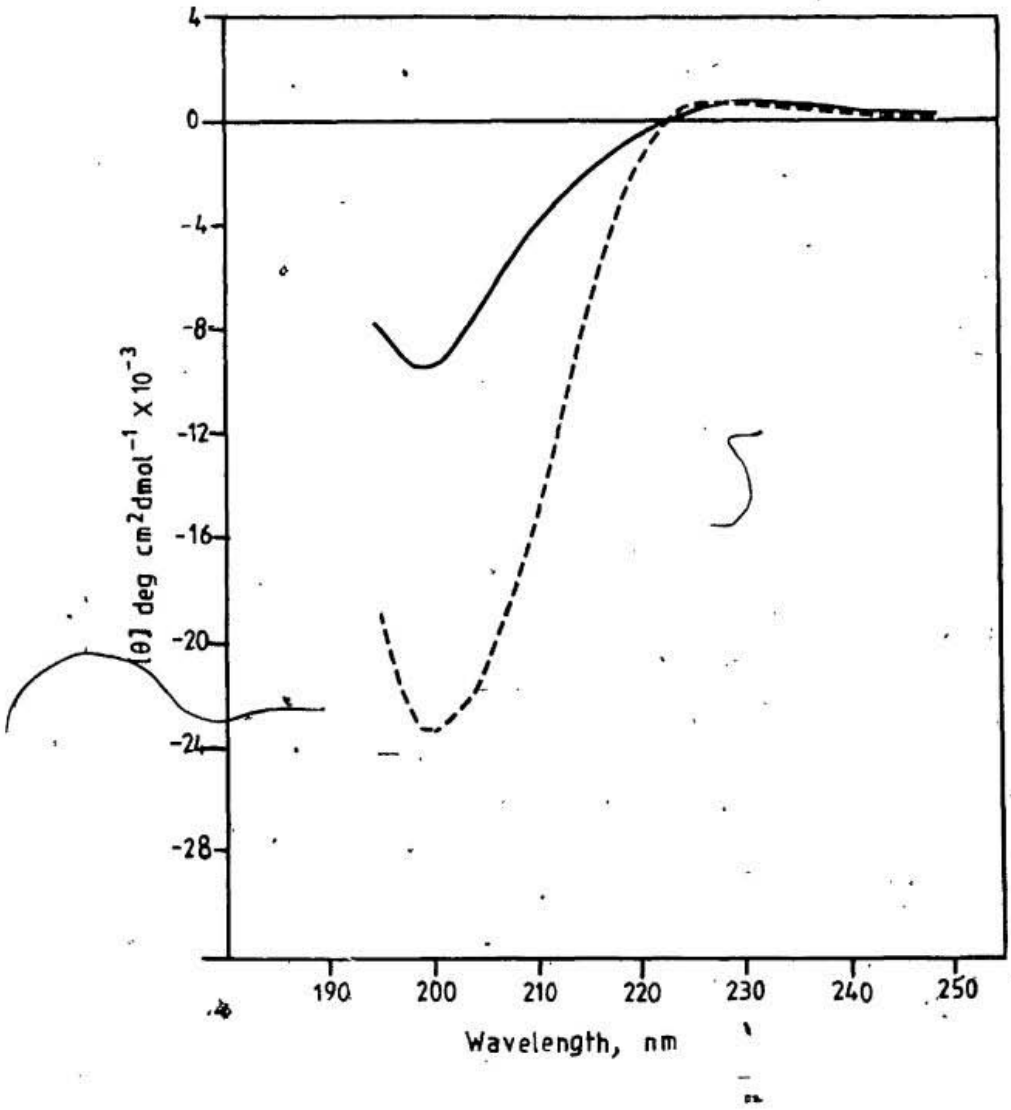


Figure A-15: CD Analysis of Boc-PPGP-$NHCH_3$ in Water

(—) Experimental spectrum; (- - -) Computed spectrum for

35% β-turn + 45% PP-II + 20% random-coil structures.

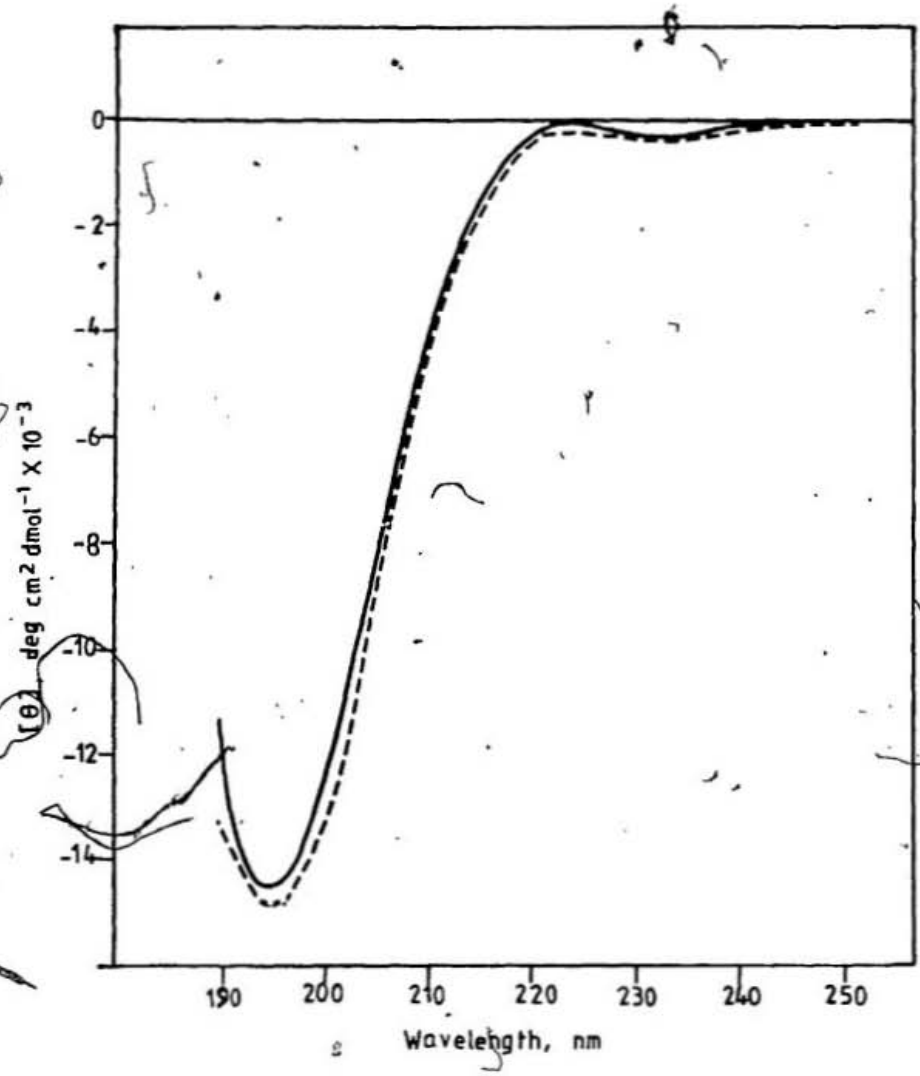


Figure A-16: CD Analysis of Boc-PPGPP-OH in TFE

(—) Experimental spectrum; (- - -) Computed spectrum for 50% β-turn + 25% PP-II + 25% random-coil structures.

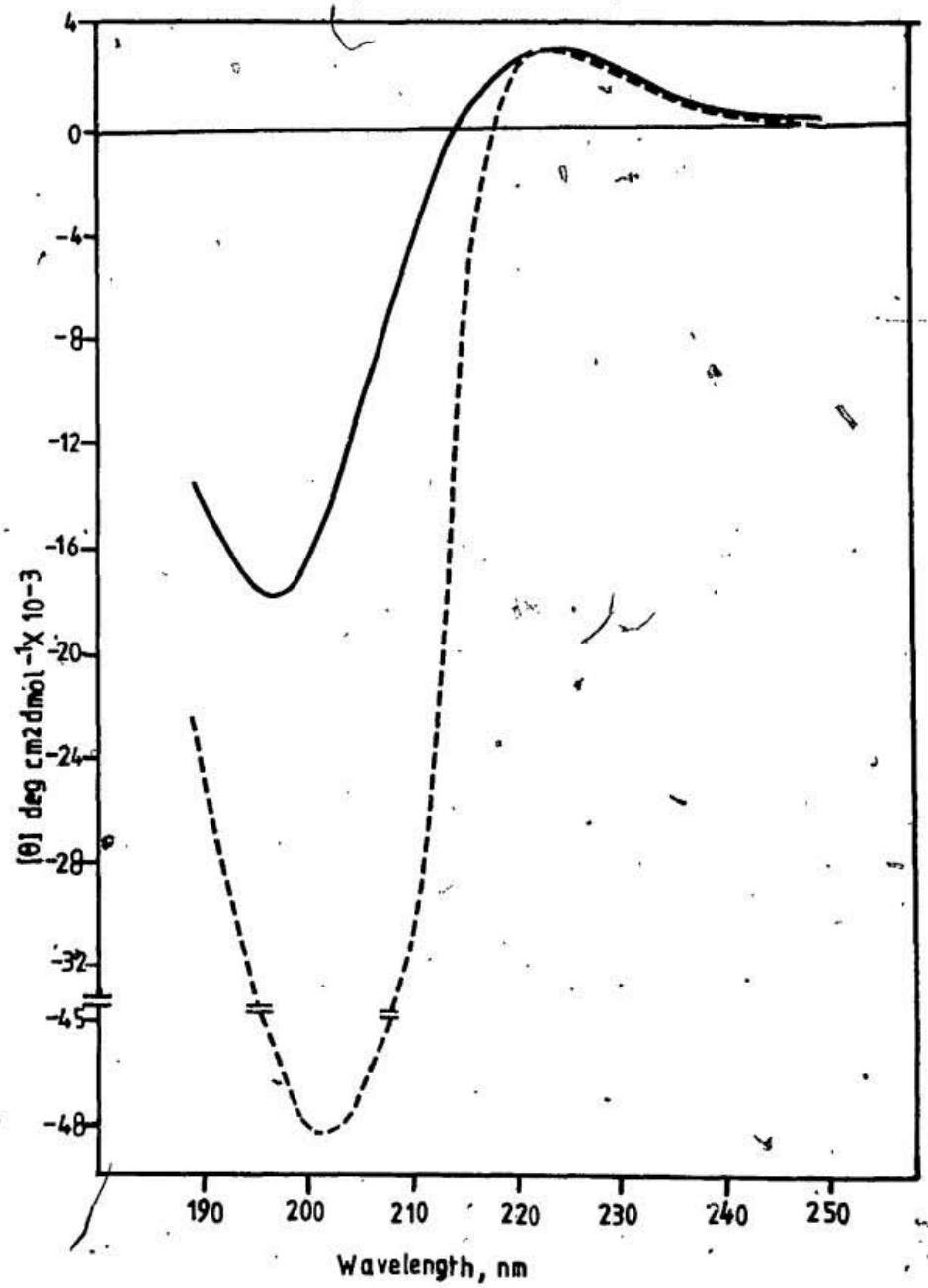


Figure A-17: CD Analysis of Boc-PPGPP-OH in Water

(—) Experimental spectrum; (- - -) Computed spectrum for

10% β-turn + 85% PP-II + 5% random-coil structures.

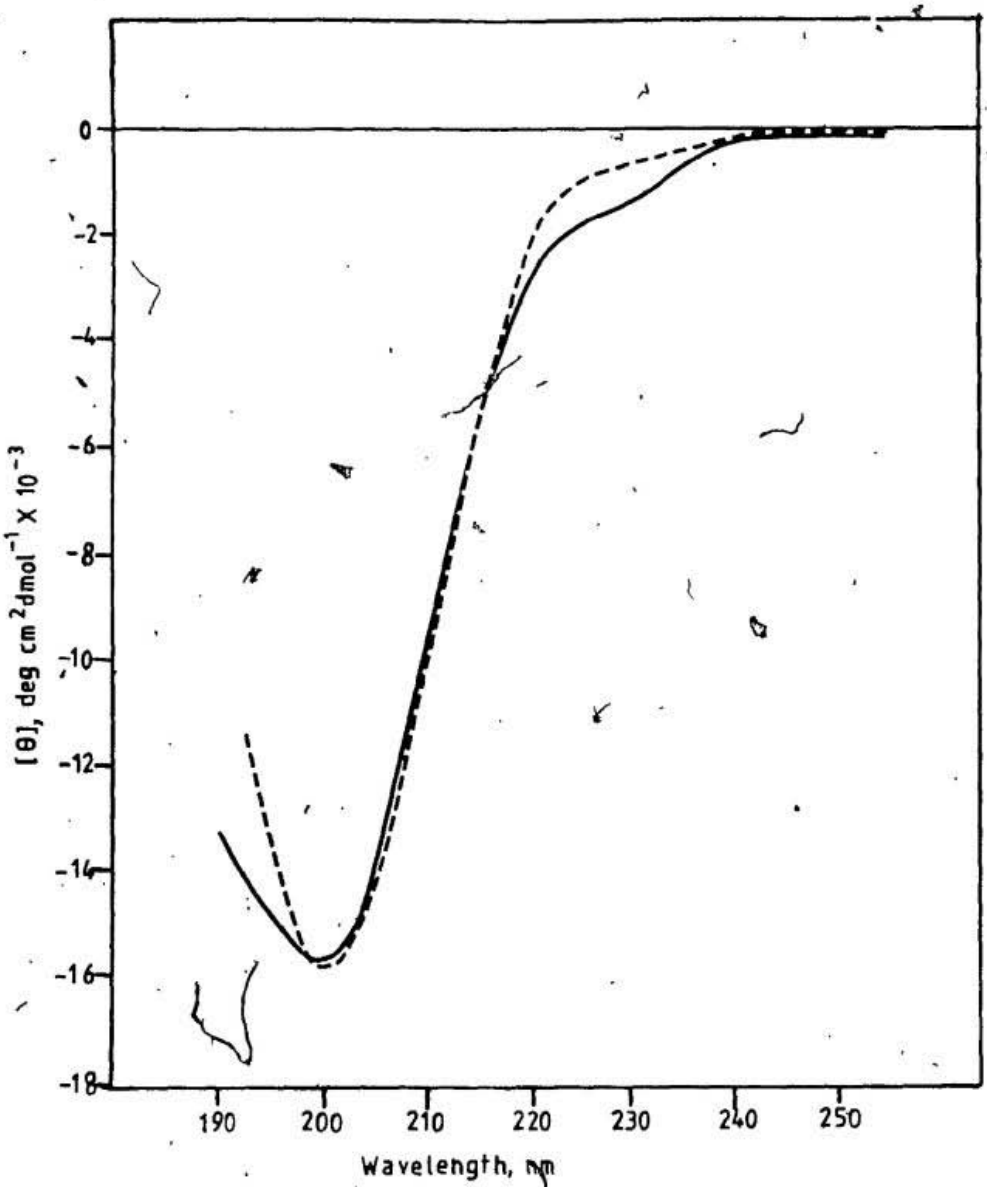


Figure A-18: CD Analysis of Boc-PPAP-OH in TFE

(—) Experimental spectrum; (- - -) Computed spectrum for 40% β-turn + 20% PP-II + 40% random-coil structures.

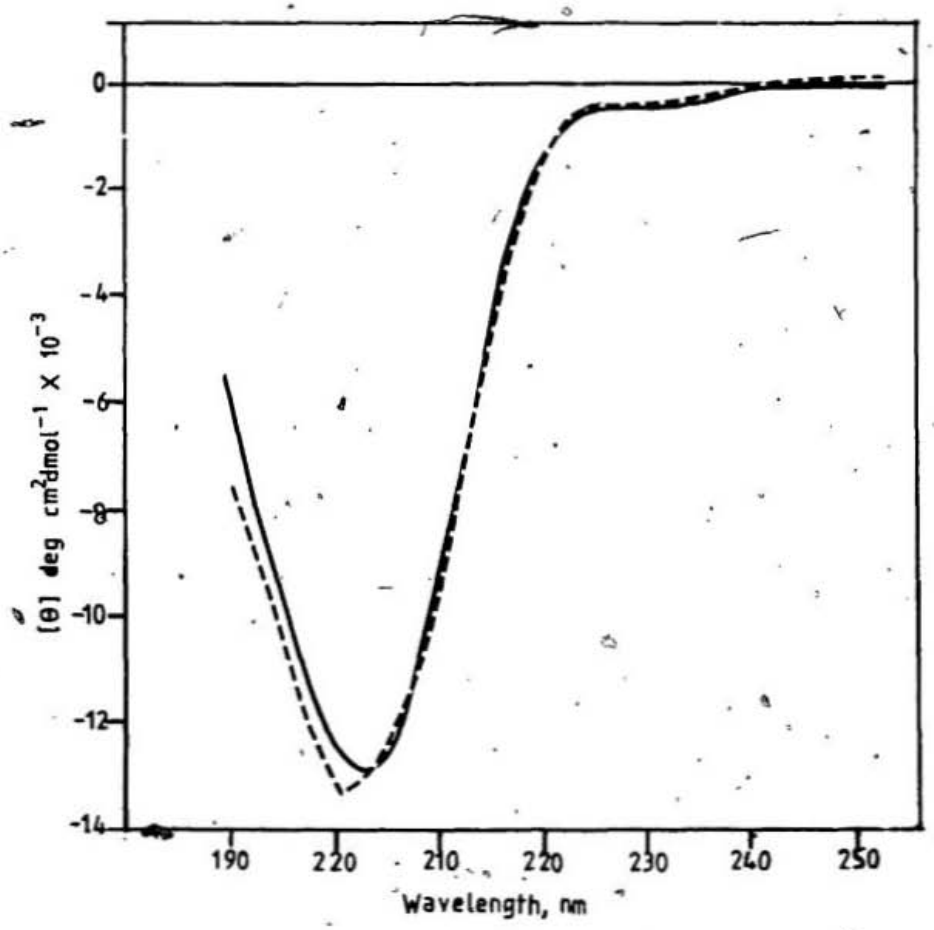


Figure A-19: CD Analysis of Boc-PPQP-OMe in TFE

(—) Experimental spectrum; (- - -) Computed spectrum for

55% β-turn + 25% PP-II + 20% random-coil structures.

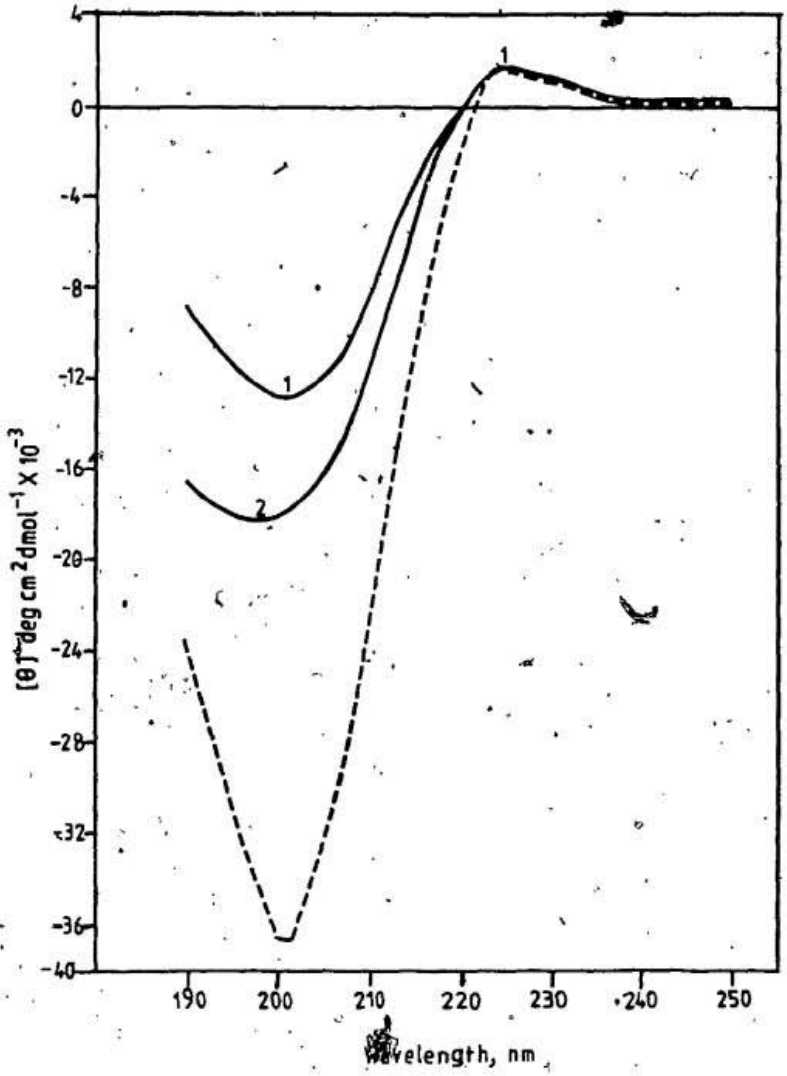


Figure A-20: CD Analysis of Boc-PPQP-OMe and Boc-PPAP-OH in Water (—) Experimental spectra of (1) Boc-PPQP-OMe and (2) Boc-PPAP-OH; (- - -) Computed spectrum for 20% β-turn + 60% PP-II + 20% random-coil structures.

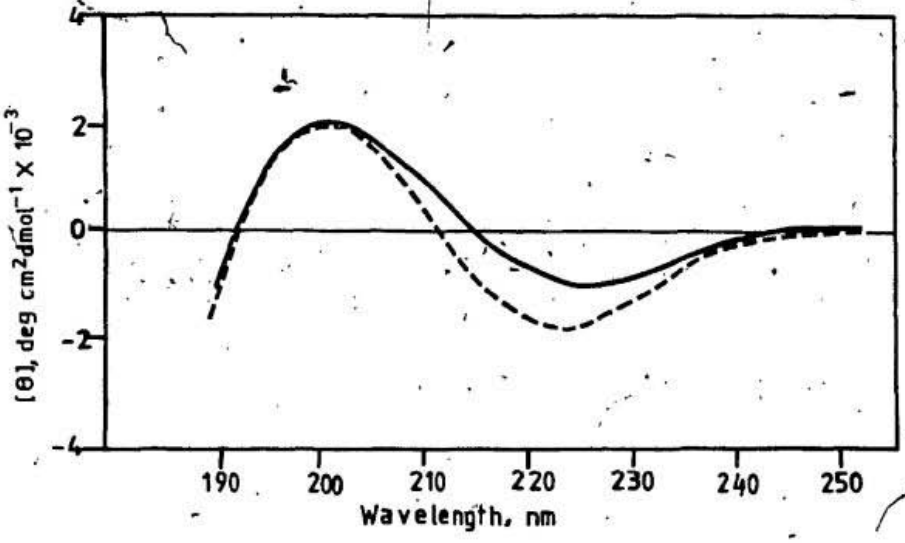


Figure A-21: CD Analysis of Boc-GVPGV-OH in TFE

(—) Experimental spectrum; (- - -) Computed spectrum for 77% β-turn + 23% random-coil structures.

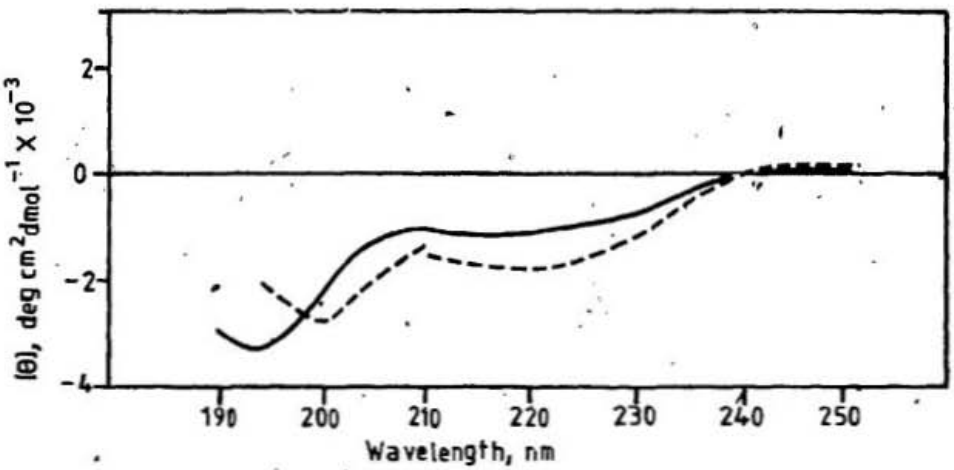


Figure A-22: CD Analysis of Boc-GVPGV-OH in Water

(—) Experimental spectrum; (- - -) Computed spectrum for

60% β-turn + 40% random-coil structures.

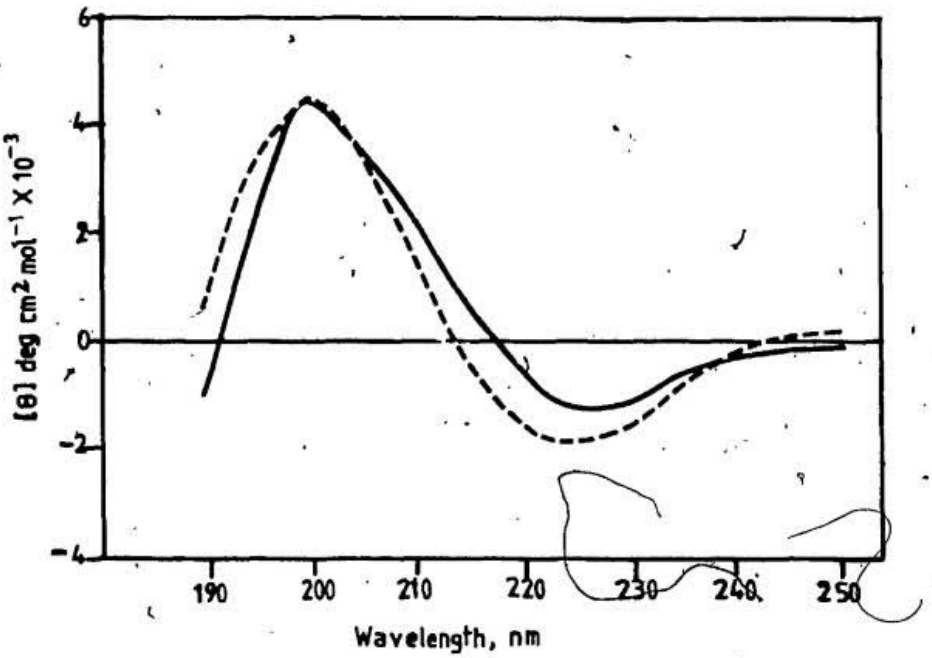


Figure A-23: CD Analysis of Boc-VPGV-OH in TFE

(—) Experimental spectrum; (- - -) Computed spectrum for 85% β-turn + 15% random-coil structures.

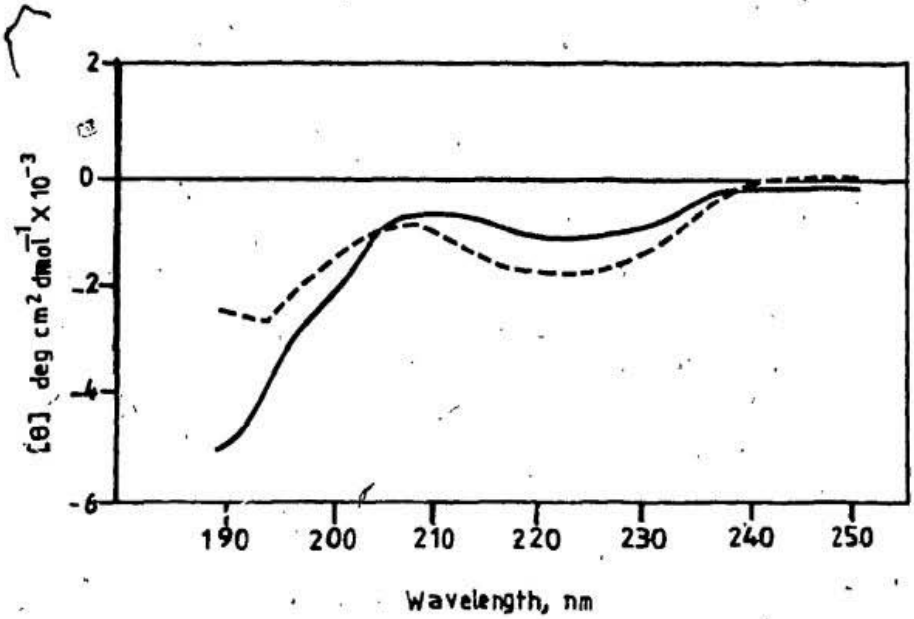


Figure A-24: CD Analysis of Boc-VPGV-OH in Water

(—) Experimental spectrum; (- - -) Computed spectrum for

65% β-turn + 35% random-coil structures.

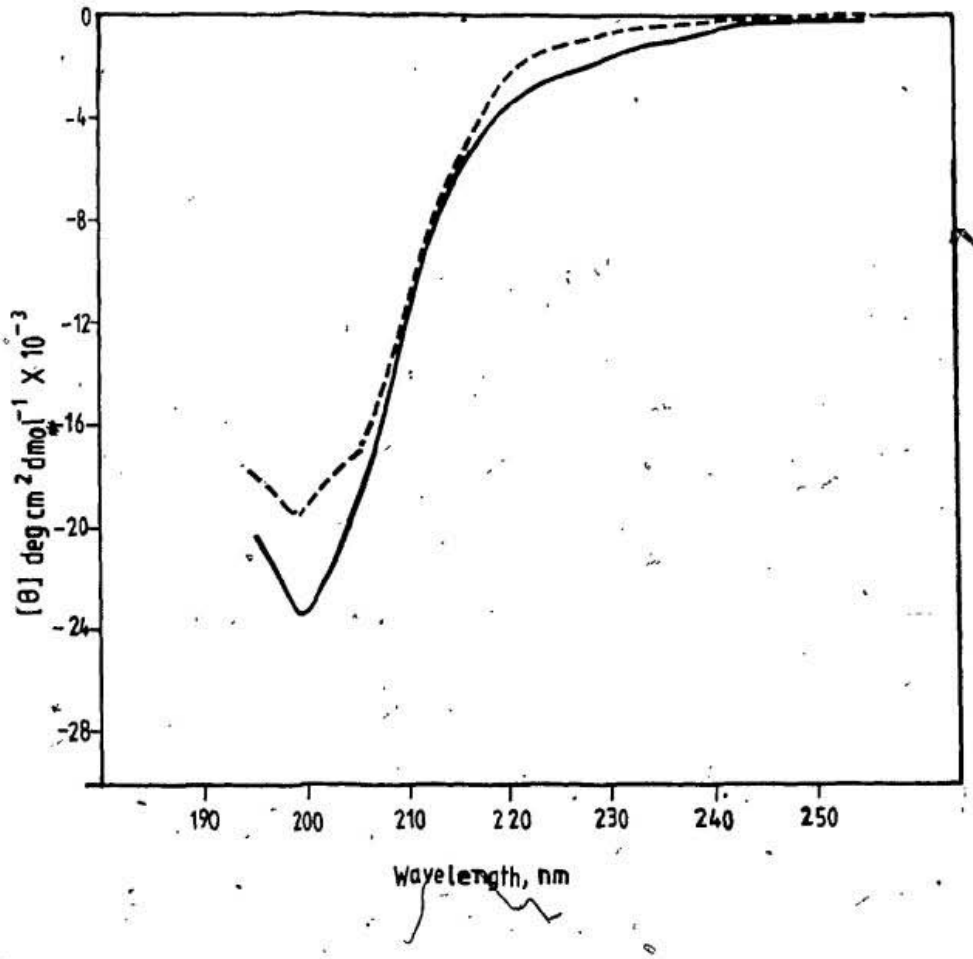


Figure A-25: CD Analysis of $(Pro\text{-}Pro\text{-}Gly)_5$ in TFE

(—) Experimental spectrum; (- - -) Computed spectrum for

40% β-turn + 20% PP-II + 40% random-coil structures.

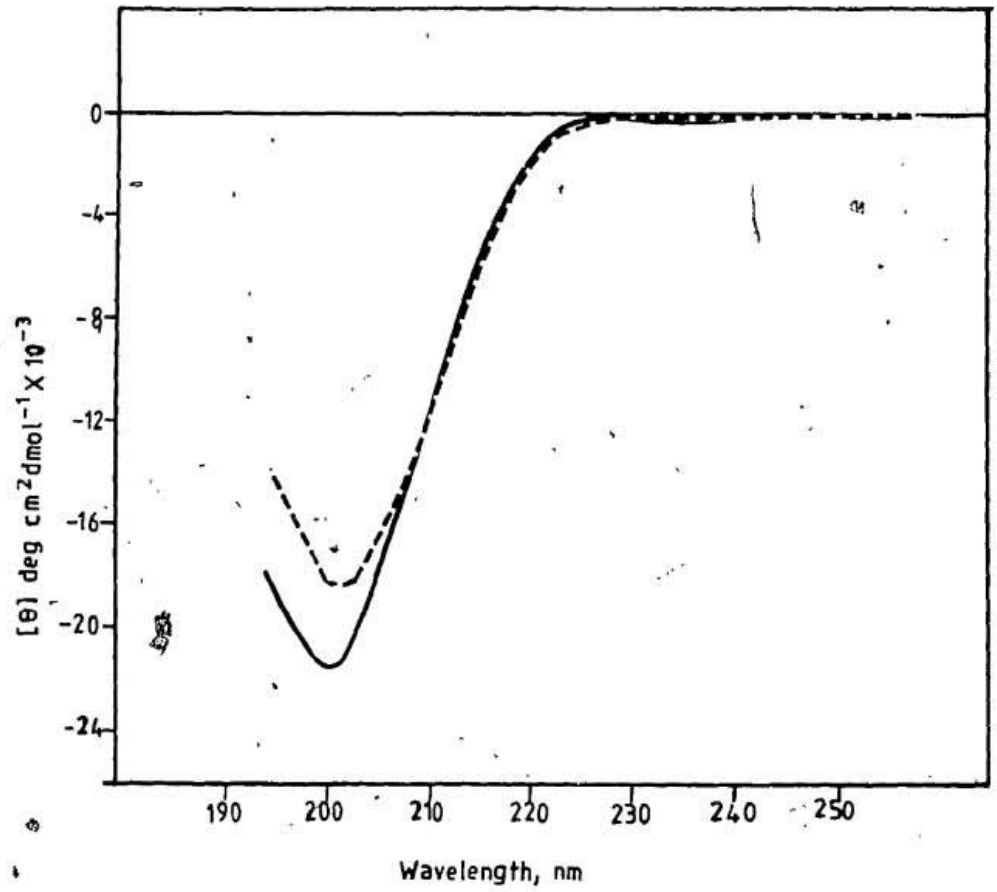


Figure A-26: CD Analysis of $(Pro\text{-}Pro\text{-}Gly)_5$ in Water

(—) Experimental spectrum; (- - -) Computed spectrum

for 35% β-turn + 25% PP-II + 40% random-coil structures.